UNCLASSIFIED

DRES

SUFFIELD REPORT

NO. 282

DEVELOPMENTS IN QUANTITATIVE STRUCTURE-ACTIVITY RELATIONSHIPS (QSAR)

A REVIEW

by

H.L. Holmes

PROJECT NO. 20-03-05

July 1976

DEFENCE RESEARCH ESTABLISHMENT SUFFIELD : RALSTON : ALBERTA

UNCLASSIFIED

DEFENCE RESEARCH ESTABLISHMENT SUFFIELD
RALSTON ALBERTA

SUFFIELD REPORT NO. 282

(6) DEVELOPMENTS IN QUANTITATIVE STRUCTURE-ACTIVITY RELATIONSHIPS (QSAR). A REVIEW

by

(10) H.L. Holmes

(14) DRES-282

PROJECT NO. 20-03-05

(11) July 1976

(12) 37 p.

403104

4/B

DEVELOPMENTS IN QUANTITATIVE STRUCTURE-ACTIVITY RELATIONSHIPS (QSAR)

A REVIEW

CONTENTS

DEFENCE RESEARCH ESTABLISHMENT SUFFIELD
RALSTON ALBERTA

SUFFIELD REPORT NO. 282

DEVELOPMENTS IN QUANTITATIVE STRUCTURE-ACTIVITY RELATIONSHIPS (QSAR)

A REVIEW

by

H. L. Holmes

ABSTRACT

Three different approaches have been developed by groups headed by Hansch, Holmes and by Free and Wilson to the problem of quantitative structure-activity relationships (QSAR) of drugs to biological processes and these have been reviewed and the strengths and weaknesses of each method annotated.

The five general equations developed by Hansch from extra-thermodynamic considerations are presented as well as the ones supplemented by inclusion of steric substituent constants and dummy or indicator parameters. Examples of 10 representative type equations are presented for the calculation of biological activities of 23 series of drugs upon 21 biological systems. More than 2000 biomedical QSAR have been developed involving more than 20,000 compounds and these are stored in Hansch's data bank. The strength of this approach lies in the versatility of these substituent constants which apply to all series of drugs. The importance of the ideal partition coefficient, P_0, and of the intercept in equations in the design of new drugs has been demonstrated. This method is rapid but in many cases only one series of drugs can be considered at one time. Furthermore wastage by metabolism and elimination are not accommodated by these equations, although equations involving the same substituent constants have been developed separately for these processes. No

combination of electronic and hydrophobic substituent constants has yet been found to adequately predict the extent of hydrogen bonding and its effect upon the activity of drugs.

The method developed at DRES is based upon evaluation of the three dominant factors (1. rate of penetration to the site, 2. rate of reaction at the site and 3. rate of wastage) governing the degrees of biological activities of drugs by *in vitro* physico-chemical properties derived from suitably chosen model experiments. By applying the same drugs to a number of different biological systems the significance of the coefficient of the partition coefficient term in these equations became apparent. Mathematical manipulation of equations derived for the same combination of drug families on two different organisms permits the calculation of the biological activities of these drugs on one organism from the observed activities upon another organism. These equations include a term to accommodate wastage,and the experimental determination of *in vitro* physico-chemical properties obviates the necessity of calculating the effects of hydrogen bonding. This method, involving the determination of *in vitro* physico-chemical properties, has advantages but the method is more time-consuming.

The Free-Wilson method involves the development of substituent constants in biological activity units from a large series of drugs in one family for one specific biological process. Addition of the cogent substituent constants to the experimentally determined biological activity of the parent drug gives the biological activity of the derivatives. These substituent constants are all inclusive,incorporating into one substituent constant the effects of metabolism, elimination and hydrogen bonding etc. The disadvantage of this method is that a new series of substituent constants must be developed for each series of drugs applied to each biological process, so it is very time-consuming and lacks the versatility of the first method.

In all some 297 equations and 207 references are presented in this review to illustrate the ways that these methods have been used to examine the mechanisms of biological processes and how this information can be used in the systematic design of more effective drugs.

SYMBOLS AND ABBREVIATIONS

The symbols used by Holmes for partition coefficients in the systems cyclohexane-water and in 1-octanol-water are the reverse of those used by Hansch. The same is true for π and π'. For consistency in this article the following symbols will be used.

Partition System	Symbols used in this Paper	Symbols used by Hansch	Symbols used by Holmes
1-Octanol-Water	P	P	P'
	π	π	π'
Cyclohexane-Water	P'	-	P
	π'	-	π

A is the agonist in moles per litre (84 page 351) or per kg of animal weight (84 page 1405).

A_{analg} is the analgesic activity of the agonist in moles per kg of animal weight (84 page 1405).

A_B is the stimulatory activity of the agonist in moles per litre in the blepharospasm test (84 page 1444).

$A_{gen\ depr}$ is the general depressant activity of the agonist in moles per kg of animal weight (84 page 1405).

$A_{resp\ depr}$ is the respiratory depressant activity of the agonist in moles per kg of animal (84 page 1405).

A_T is the threshold stimulatory activity in moles per litre on the frog flexor reflex (84 page 351).

BR is biological response usually recorded as $\frac{1}{C}$.

C is concentration in moles per litre.

Ca is cats.

D is dummy or indicator parameter (60).

$E_{\frac{1}{2}}$ is the polarographic half-wave potential against a saturated calomel electrode (84 page 1210).

E.c. is *Escherichia coli*.

E_R is a homolytic equivalent of σ for free radical reactions (110).

E_S is the Taft steric factor (27,44).

I is the indicator or dummy parameter (60).

IG_{50}^{17} is the 50% inhibition of growth caused by the compound expressed in moles per litre after 17 hours incubation (84 page 1344).

k_{SH} is the second order rate constant for the addition of n-butanethiol to conjugated heteroenoid compounds (84 page 1105).

k_W is the pseudo first order rate constant for the hydrolysis of conjugated heteroenoid compounds by a reverse aldol process (84 page 1104).

k_X is the rate constant for some reaction with an aromatic system bearing a substituent, X, on the benzene ring and for which k_H is the rate constant for that reaction on the parent compound of that family.

K_{SH} is the equilibrium constant for the reversible addition of n-C_4H_9SH to the conjugated heteroenoid compounds (84 page 213).

Log P is the logarithm of the partition coefficient of the solute in the system 1-octanol-water (4). In Holmes' work (84) this is log P': see the first item.

Log P' is the logarithm of the partition coefficient of the solute in the system cyclohexane-water. In Holmes' work (84) this is log P: see the first item.

Log P" is the logarithm of the partition coefficient of the solute in the system cyclohexanol-water (84 page 776).

Log P'" is the logarithm of the partition coefficient of the solute in the system ether-water (84 pages 777 and 781).

Log P_0 is the ideal or maximum value of log P for an agonist to cause maximum biological response (41). This is derived from equations involving $(\log P)^2$ and log P by taking the partial derivative of log BR with respect to log P and setting the partial derivative equal to zero and solving for log P. This value is log P_0.

M is mice.

MR is substituent molar refraction (92).

μ is dipole moment.

n is the number of compounds considered in a multiple regression analysis.

π is the increment (in log units) in log P due to a substituent in the system 1-octanol-water (4)(see π' reference 84 page 795).

π_0 is the substituent equivalent of log P_0 (3).

π' is the increment (in log units) in log P' due to a substituent in the system cyclohexane-water (84 page 787 see π in Holmes' nomenclature).

$\pi'_{interaction}$ is the increment (in log units) in log P' in the system cyclohexane-water due to interaction between contiguous groups. (84 page 787).

r is the correlation coefficient for the results derived from equations developed by a multiple regression analysis of biological and physicochemical data.

R is rats.

Ra is rabbits.

rho, ρ, is the reaction constant in Hammett's equation log k_X - log k_H = $\rho\sigma$. It is a constant for all substituents and depends only on the reaction series (90,91).

s is the standard deviation.

S.a. is *Staphylococcus aureus*.

S.alb. is *Staphylococcus albus*.

σ is the substituent constant in Hammett's equation $\log k_X - \log k_H = \rho\sigma$. It is determined by the nature of the substituent and independent of the reaction, the constant k_X of which is involved in the equation (90,91,92,103).

σ^+ is Brown's σ constant (105).

σ^- is Hammett constant for substituted phenols and amines, etc. (44,103).

σ^* is Taft's σ constant for aliphatic compounds (106,107).

$\sigma\cdot$ is the substituent constant for homolytic or free radical reactions (116,117).

$\mathcal{F}$ is Swain and Lupton's inductive component in Hammett's σ (108).

$\mathcal{R}$ is Swain and Lupton's resonance component in Hammett's σ (108).

DEFENCE RESEARCH ESTABLISHMENT SUFFIELD
RALSTON ALBERTA

SUFFIELD REPORT NO. 282

DEVELOPMENTS IN QUANTITATIVE STRUCTURE-ACTIVITY RELATIONSHIPS (QSAR)

A REVIEW

by

H. L. Holmes

INTRODUCTION

Search for a new drug can be approached in two ways. This can be done by the random synthesis and biological testing of new compounds or it can be approached systematically using the knowledge that has accumulated from structure-activity studies. Spinks (85)* has examined the problem of finding new drugs and from his analysis he estimates that one new drug arises out of each 200,000 compounds. He also indicates that one can expect to find an anticancer drug out of each 400,000,000 randomly tested compounds. This figure is about 100 times the number of known organic compounds. On the other hand, knowledge of the factors governing the relationship between structure and biological activity is growing by leaps and bounds (71,72). Hansch (70) has more than 2000 biomedical QSAR on more than 20,000 organic compounds in his data bank alone. The group at DRES applied the same 900 compounds to a large number of biological systems and as a result over 500 QSAR on some 900 compounds are similarly stored. In a search for a new drug then, one is faced with the choice of hiring a very large staff of synthetic organic chemists for a long period of time, to say nothing of the large biological test team necessary, or staking one's chances upon a QSAR approach or upon an intelligent combination of both.

References 1-84 cover quite completely the QSAR work of Hansch, Holmes, Lien, Fujito and their collaborators.

The embryonic concept which eventually bloomed into the structure-activity relationship dates back to 1870 when Crum-Brown and Fraser (86) sensed that biological response, BR, to drugs was related to their chemical structures, CS. This relationship may be expressed mathematically by equation 1.

$$BR = f(CS) \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad 1$$

They stated that it should be possible to develop a calculus of SAR by making small changes in chemical structure and relating these to BR. The obstacle that prevented them from realizing their dream was the problem of defining significant parameters of chemical structure in numerical terms.

Two developments in the first third of this century laid the foundation for the development of numerical values for evaluating "chemical structure". At the turn of the century Meyer (87) and Overton (88) used oil-water partition coefficients to quantitatively evaluate the penetration of simple organic compounds which act as anaesthetics. It was argued that compounds with high oil-water partition coefficients should be good anaesthetics. While this generalization was true for a number of compounds, it was not universally true. Other factors must also be operating in the control of anaesthetic activity and these were not apparent to Meyer and Overton so they were unable to formulate an equation relating biological response, BR, to various factors evaluating chemical structure, CS. A third of a century of qualitative study of the electronic effects of substituents by the Lapworth-Robinson School culminated in 1935 in Hammett's formulation of numerical constants, σ, for electronic effects of substituents (89,90). This simple but extremely important idea opened the floodgate for the proliferation of specialized substituent constants, $\sigma^*\sigma^+$, σ^- and $\sigma^\cdot$ and this led to the resolution of σ into inductive and resonance components (91). Steric constants E_S were then developed by Taft (91). Molecular refraction (92) also provides, besides a measure of polarizability, a measure of bulk volume. Following the lead of Hammett in developing free energy related substituent constants, Hansch (4) in the early 1960's developed

a substituent constant π which is the increment (in log units) in the logarithm of the partition coefficient, P, due to a substituent, X. This may be expressed as in equation 2. Log P and π adequately evaluated the "random walk" of

$$\log P_X - \log P_H = \pi \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad 2$$

drugs to the site of action in biological systems and also the hydrophobic binding of organic compounds to enzymes, serums and mitochondrial proteins (6,20). This permitted the expansion of equation 1 to 3.

$$\Delta BR = f(\Delta \text{ hydrophobic} + \Delta \text{ electronic} + \Delta \text{ steric} + \Delta \text{ polarizability}). \quad 3$$

From extra-thermodynamic considerations Hansch developed the general equations 4 and 5 relating biological response to substituent constants and molar refraction. The k_1, k_2 etc. of equation 4

$$\log BR = -k_1 (\log P)^2 + k_2 \log P + k_3\sigma + k_4 E_S + k_5 MR + k_6 \quad \ldots\ldots \quad 4$$

$$\log BR = -m_1\pi^2 + m_2\pi + m_3\sigma + m_4 E_S + m_5 MR + m_6 \quad \ldots\ldots\ldots\ldots \quad 5$$

and the m_1, m_2 etc. of equation 5 are constants. Under specified conditions (3) these equations may be reduced to ones with fewer terms. In the extra-thermodynamic development of equations 4 and 5 it was tacitly assumed that the rate of wastage (metabolism and elimination) of the drug in the biological test system is either zero or a constant value for the family of drugs under consideration. The word family is used because incorporation of Hammett's sigma, σ, constant limits the equation to the examination of one family of drugs at a time. Equations have been set up by Hansch (18) relating 1) the biological response BR to a given set of parameters and 2) the rate of wastage to the same set of parameters, which provides an insight into which member of the family provides the happy compromise between the degree of biological response and the rate of wastage.

The advent of the computer enabled Hansch, Holmes and others to examine the action of drugs on many biological systems and to determine the dominant factors governing the degree of the biological response. Developing equations by regression analysis for one, two, three, etc. terms on the right hand side of equations 4 or 5 gave correlation coefficient, r, for the value of BR, calculated from the equation compared to the observed value, as well as the standard deviation, s. The equation with the largest correlation

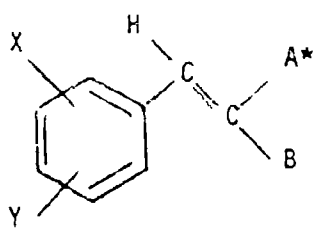
coefficient and the smallest standard deviation would be the equation best fitting the data. The terms on the right hand side of the equation reveal the dominant factors governing the degree of the biological response. If the coefficient of the E_S term is zero, then steric factors do not play a role in this reaction. If log BR for the antifungal activity of aniline-phenol mixtures against *Candida albicans* is measured as log $\frac{1}{C}$ where C is the molar concentration necessary to produce the biological action, then log $\frac{1}{C}$ is given by equation 6 (35). Obviously the rate of penetration (log P) and the rate of reaction (σ) at the site

$$\log \frac{1}{C} = 0.555 \log P + 2.193 \sigma - 1.322$$

$$r = 0.982, s = 0.111 \quad \ldots\ldots\ldots\ldots \quad 6$$

of action are the dominant factors governing log $\frac{1}{C}$ in equation 6.

Holmes (84), in his approach, introduced a term for the rate of wastage in the biological process but has not yet extended his equation to include terms for steric factors or polarizability. To evaluate the rates of reaction at the site of action he used *in vitro* rate constants, k_{SH}, for the addition of n-C_4H_9SH instead of substituent constants. For the bacterial growth inhibitory activities, IG_{50}^{17}, of the conjugated heteroenoid compounds, I(A = $COCH_3$, $CO_2C_2H_5$, $CONH_2$; B = $COCH_3$, $CO_2C_2H_5$, $CONH_2$), the *in vivo* rates of wastage were evaluated by



I

---

* In this generalized formula A will always be considered to be *trans* to the phenyl group and B *cis* to it.

the *in vitro* rates of hydrolysis, k_W, of the I compounds by reverse aldol process. Multiple linear regression analysis of the data for 52 compounds gave equation 7.

$$\log IG_{50}^{17} = -0.24 \log P'^{*} - 0.55 \log k_{SH} + 1.07 \log k_W + 0.72 \quad . \quad . \quad . \quad 7$$

This 50% inhibition of growth, IG_{50}^{17}, of *Staphylococcus aureus* is for 17 hours at 37°C. Determination of the stimulatory activities, A_T (in moles per litre), against the frog flexor reflex was complete in 5 minutes so wastage was not a significant factor and the coefficient of $\log k_W$ in equation 8 is zero. The correlation coefficient for equation 7 was good so $\log IG_{50}^{17}$ (calc)

$$\log A_T^{**} = -0.21 \log P' - 0.41 \log k_{SH} - 4.59 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad 7a$$

from equation 7 is about equal to $\log IG_{50}^{17}$ (obs). Subtraction of equation 7 from equation 7a and replacement of $\log IG_{50}^{17}$ (calc) by $\log IG_{50}^{17}$ (obs) gives equation 7b. This and similar equations permits the calculation of the

$$\log A_T = \log IG_{50}^{17} \text{ (obs)} + 0.03 \log P' + 0.14 \log k_{SH} - 1.07 \log k_W - 5.31 \quad . \quad . \quad . \quad . \quad 7b$$

biological activities of a family of drugs on one organism from the observed activities on another organism.

An alternative method for relating the degrees of biological activity of a family of drugs to *de novo* substituent constants was introduced by Bruice *et al.* (93) and more fully developed by Free and Wilson (94). Free and Wilson have developed a set of substituent constants for a series of ten tetracyclines involving different substituents in three positions on the parent ring system. By setting up a series of simultaneous equations, one for

---

* To avoid confusion between the terminology of Hansch and Holmes, the following symbols have been adopted throughout this review. P is the partition coefficient in 1-octanol-water, P' is the partition coefficient in the system cyclohexane-water and P''' is used for the system ether-water. This means that P and P' are reversed in the article by Holmes (84).

** $\log \frac{1}{C}$ of equation 6 equals $\log 1 - \log C$. This is equal to $-\log C$. Log A_T then will be comparable to $-\log \frac{1}{C}$.

each substituent in each position, it was possible to assign substituent constants (in biological units) so that the biological activity is the sum of these biological substituent constants plus a constant. This method lumps into one constant hydrophobic effects of the substituent along with electronic, steric factors and group interactions. As well metabolism and elimination are accommodated in these constants. The disadvantage of this approach is that constants derived for one set of drugs will not be useful for another set causing a different biological response. This is the advantage of the free energy related substituent constants developed by Hammett, Taft and Hansch etc. They have been successfully applied to over 2000 biomedical QSAR.

Strategy in drug modification is becoming a subject in its own right (95). Techniques, independent of the computer, have been developed by Topliss (41,96,97) and Craig (98) for arriving at the derivative in a congeneric series of drugs with the optimum properties, by examination of a minimum number of members of the series. Cluster analysis (99) is being considered as a means of minimizing the problem of collinearity between terms in multiterm equations. As our understanding of drug action at the molecular level advances through the development of biochemistry, enzymology (100-102) and molecular biology it becomes more urgent to systematize the interactions of organic compounds with as many biochemical systems as possible. The explosion of data pouring from the current journals is a waste of effort unless it can be retrieved and correlated with other data. This is one of the pressing problems in the field of Structure-Activity at the moment. The method of data management being developed by Hansch is reported in three of his recent articles (70-72).

These various approaches will be discussed in detail as well as the method for deriving the ideal log P value, log P_0, for the member of a series of drugs when applied to some biological system. This will be used in the systematic approach to the design of more effective drugs.

FREE ENERGY RELATED SUBSTITUENT CONSTANTS (92)

Substituent constants σ and E_S have been developed for homogeneous reactions but, surprisingly, experience has shown that these apply equally as well to heterogeneous reactions of biological systems. For example E_S adequately accommodates the steric interactions between substrate and biological system (40). Log P or π fulfill a non-specific role and a specific role (37). By its non-specific role it evaluates the "random walk" of the drug to the site of action, while the specific role evaluates the binding of organic compounds to the enzyme (13). These constants will be discussed in general terms for those not mathematically oriented.

Electronic Substituent Constants (90,91,92,103)

Determination of rates of reaction between organic compounds or substrates and the biophase at the site of action in a biological system would be difficult or impossible. However the product of the Hammett reaction constant ρ and the substituent constant σ has enabled Hansch to obviate this difficulty in his approach to the Structure-Activity problem and he has exploited it to the full.

To dispel any confusion as to the meaning of electronic as employed in this context it will be considered to mean the following. When a drug reacts with a receptor site with the formation of a covalent bond, the reaction will be very sensitive to quite small changes in the electron density on and around the atom involved in the reaction. Such electronic differences could easily be small enough to have little or no effect upon other parameters, such as partition coefficients, but still have a large effect upon the covalent bond forming reaction. Craig (98) has examined the interdependence of the variables used in structure-activity relationships and Hansch (44) has shown in equations 8 and 9 that there is a slight degree of covariance between π and σ. Equation 8 relates π values for meta substituents to σ_m for similarly located groups while equation 9 does the same for para substituted groups.

$$\pi = -1.84\ \sigma_m + 0.70$$

$$n = 34,\ r = 0.387,\ s = 1.079 \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad 8$$

$$\pi = -0.89\ \sigma_p + 0.48$$

$$n = 37,\ r = 0.300,\ s = 1.110 \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad 9$$

The Hammett reaction constant ρ and the substituent constant σ are based upon the observation of Hammett (89) that for a family of benzene derivatives there is a linear relationship between the logarithms of the rate constants, k_X, or equilibrium constants and the logarithms of the ionization constants, K_X, of the similarly substituted benzoic acids. This may be expressed mathematically as in equation 10 where the slope ρ is a measure of the sensitivity of the

$$\log k_X = \rho \log K_X + C \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad 10$$

logarithm of k_X to electronic effects of the substituent, X. K_X is the ionization constant of the similarly substituted benzoic acid. When X = H is substituted into equation 10, as in equation 11, this gives the relationship for the parent compounds of these two series.

$$\log k_H = \rho \log K_H + C \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad 11$$

Subtracting equation 11 from equation 10 gives equation 12.

$$\log \frac{k_X}{k_H} = \rho \log \frac{K_X}{K_H} \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad 12$$

Log $\frac{K_X}{K_H}$ was defined by Hammett as σ

$$\log \frac{K_X}{K_H} = \sigma \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad 13$$

and substituting σ for this term in equation 12 gives equations 14 and 15.

$$\log \frac{k_X}{k_H} = \log k_X - \log k_H = \rho\sigma \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad 14$$

$$\log k_X = \rho\sigma + \log k_H \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad 15$$

The reaction constant, ρ, by the nature of the linear relationship is a constant for all substituents and depends only upon the reaction

series being examined. The substituent constant, σ (in log units), is by definition, determined by the nature of the substituent and is independent of the reaction with which k is associated. Tables of substituent constants, σ, appear in references 90, 91 and 92. Substituent constants, σ, for di-substituted derivatives can be approximated quite well from the algebraic sum of the σ constants for each of the substituent groups. Thus the σ constant for 3,4-dimethoxy derivatives is the algebraic sum of σ_m for the methoxyl group + σ_p for the methoxyl group. While equation 15 gives amazingly good correlations for a wide variety of reactions involving *meta*- and *para*-

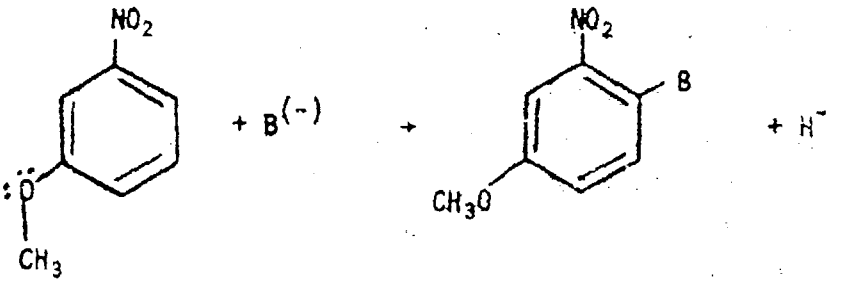
$$\sigma_{3,4\text{-di-}OCH_3} = +0.115 + (-0.268) = -0.153 \quad \ldots\ldots\ldots\ldots \quad 16$$

substituted benzene derivatives (see ref. 84 pages 215, 220-224,228,236, 243 and 302), it has its limitations. In the case of *ortho*-substituted benzene derivatives, steric factors are superimposed upon electronic effects so a different set of constants (104) must be used. Equation 15 does not hold for aliphatic compounds.

There are two types of benzene derivatives for which Hammett σ constants fail to provide reliable results. The first one is where a substituent containing lone electron pairs is coupled through a series of benzene double bonds (?) to an electron-withdrawing or accepting group (e.g. CHO, CN, NO_2 etc.) as in II and III, then σ^- constants (44) should be used. A second situation where σ constants fail is when an electron-donating substituent is so oriented as to reduce the positive charge at a centre on the benzene ring involved in electrophilic substitution as in IV → V.

R = H, C_nH_{2n+1}, COR

II III



IV V

Under these circumstances, σ^+ constants developed by Brown (105) and listed in references 44 and 105 must be used.

Taft *et al* (106,107) attempted to resolve σ values into their inductive and resonance components σ_I and σ_R while Swain and Lupton (108) developed the analogous constants $\mathcal{F}$ and $\mathcal{R}$. Tables of these are listed in ref. 9

Unlike the planar aromatic derivatives with *meta* and *para* substituents plots of log k values for aliphatic reactions against log K for ionization constants of aliphatic acids usually are not linear (see equation 11). Steric effects and field effects in the latter case are not sensibly constant giving rise to a scatter of points in the last plot. Taft (91), from a determination of the rates of base and acid catalyzed hydrolysis of esters of aliphatic acids of the type of X - $CH_2CO_2C_2H_5$ developed a σ^* constant and a steric constant, E_S*. Tables of σ^* values are listed in references 44 and 91, while E_S values are to be found in references 44 and 27. The greater the steric effect of a group the larger E_S is in a negative sense (66,68).

Two sets of analogous constants σ' (21,109,116,117) and E_R (110) have been developed for homolytic or free radical reactions.

---

* E_S values apply equally well to aliphatic and aromatic substituents (44).

Hydrophobic Substituent Constants, π and π'.

The free energy related ρ and σ constants of Hammett prompted Hansch to develop an analogous hydrophobic substituent constant, π*, from partition coefficients in 1-octanol-water, while Currie *et al* (111-114,84) followed with a set of π'* and $\pi'_{interaction}$ constants for the systems cyclohexane-water and 1-octanol-water.

Many partition systems have been used (115) in the past but recently partitioning has centred around the 1-octanol-water (4,84) and the cyclohexane-water systems (84). Comparing results for the two systems reveals that, over the central part of the spectrum of solutes examined, a linear relationship exists but divergence in opposite directions occurs at the two extremes. This relationship is expressed in equations 17-20 (84). Similar relationships

$$\log P = 1.85 \log P' - 2.53$$
$$n = 91,\ r = 0.91 \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad 17$$

$$\log P = -0.28 (\log P')^2 + 2.98 \log P' - 3.50$$
$$n = 91,\ r = 0.93 \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad 18$$

$$\log P = -0.088 (\log P')^3 + 0.23 (\log P')^2 + 2.22 \log P' - 3.34$$
$$n = 91,\ r = 0.93 \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad 19$$

$$\log P = -0.016 (\log P')^4 + 0.045 (\log P')^3 - 0.115 (\log P')^2 + 2.514 \log P' - 3.349$$
$$n = 91,\ r = 0.93 \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad 20$$

between partition coefficients in the system 1-octanol-water with cyclohexanol-water, P" (84), and with n-butanol-water, P^{IV} (44) are given respectively by equations 21 and 22.

$$\log P'' = 0.72 \log P + 0.60$$
$$n = 14,\ r = 0.97 \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad 21$$

$$\log P^{IV} = 0.70 \log P + 0.38$$
$$n = 57,\ r = 0.993,\ s = 0.123 \quad \ldots\ldots\ldots\ldots \quad 22$$

---

* To avoid confusion, log P and π refer to the system 1-octanol-water while log P' and π' refer to the system cyclohexane-water. This is just the reverse of the symbols used by Holmes in Structure-Activity Relationships (84).

It is known that, in many cases, water penetrates into biological media by pinocytosis* but most organic compounds appear to penetrate into biological media by an iter ted process of partitioning between nonpolar and polar media. In analogy with equations 17, 21 and 22 two partition systems could be envisaged such as

System No. 1	System No. 2
octanol + drug	Biological protein + drug
$\uparrow \downarrow K_1$	$\uparrow \downarrow K_2$
water + drug	water + drug

where equation 23 would hold. Equation 23 relates the penetration of

$$\log K_2 = a \log K_1 + b \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad 23$$

drugs into biological protein to log P (1-octanol-water). It would be interesting to test this if a biological system were available whose response to drugs was dependent only upon partition coefficients. While this is out of the question, nevertheless there are many pure enzymes whose inhibitions aredependent only upon these hydrophobic forces or partition coefficients (44,46).

By convention, the equilibrium constant, P, for a solute in the system 1-octanol-water is defined by equation 24 which holds for the process expressed in equation 25.

$$P = \frac{[\text{solute in 1-octanol}]}{[\text{solute in water}]} \quad \ldots\ldots\ldots\ldots \quad 24$$

$$\text{Solute in water} \rightleftarrows \text{Solute in 1-octanol} \quad \ldots\ldots\ldots \quad 25$$

Log P then is a measure of the free energy involved in the reversible transfer of 1 mole of solute from water to 1-octanol. Hansch (20) states that "it has been shown that the transfer of a hydrocarbon solute

---

* Pinocytosis is the penetration of compounds through holes in cell membranes (68).

from a nonpolar environment to an aqueous one is exothermic for aliphatic hydrocarbons and approximately athermal for aromatics. The solubility of these and other organic compounds in water is associated with the large negative entropy of solution which is due to the formation of loosely held but highly structured envelopes of water molecules around the apolar portions of the organic molecules as they enter the water. It is predominantly the molecular size and shape which determines how many molecules enter into the structured sheath around the apolar portions of the organic solute molecules and therefore determines the magnitude of the negative entropy of solution." Conversely the transfer of an organic solute from water to a nonpolar solvent will strip the structured sheath of water molecules from the solute molecule which may result in the generation of a weak (bond energy of about 1 k cal/mole) hydrophobic bond between the nonpolar portion of the solute and the nonpolar solvent. Hansch (20) further states that "the major factor determining the partitioning of organic molecules between aqueous and organic phases is the extent to which they form hydrophobic bonds."

Hansch developed a free energy dependent substituent constant π as the increment (in log units) in log P due to a substituent X when introduced into the parent compound, H. Equation 26 expresses this relationship. From a series of compounds values for π were deduced (92,84) which, except for minor

$$\pi = \log P_X - \log P_H \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad 26$$

variation, hold from series to series. In fact log P for a compound can be calculated from equation 27.

$$\log P = \Sigma\pi \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad 27$$

Leo and Hansch (36) developed 21 equations by regression analysis relating log P values for compounds in 2: partitioning systems to those in the system 1-octanol-water. When solutes that could hydrogen bond were combined in the series with those that could not, the correlation coefficient for the equation relating log P' to log P (analogous to equations 21 and 22) was poor. Actually two equations were needed, one to accommodate the solutes that could hydrogen bond and another for those that could not. It is claimed that solvents such as butanol, pentanol, cyclohexanol and 1-octanol*, which dissolve considerable amounts of water do not require two equations.

---

* 1-Octanol dissolves 2.3 moles of water per litre while 1-octanol is only soluble in water to the extent of 4.5×10^{-3}M (36).

Currie *et al* (111) developed an analogous series of Π' values (see reference 84 page 787) for atoms or atom groups such that log P' for a compound is given by equation 28. For 73 I compounds with B=H

$$\log P' = \Sigma\Pi' - 1.30 \quad \ldots \quad 28$$

these Π' values gave log P' (calc) values that compared well with the observed values as seen in equation 29 (84). Log P' values were then

$$\log P' \text{ (calc)} = 1.00 \log P' \text{ (obs)} + 0.04$$
$$n = 73,\ r = 1.00 \quad \ldots \quad 29$$

$$\log P' \text{ (calc)} = 1.51 \log P' \text{ (obs)} - 1.80$$
$$n = 91,\ r = 0.98 \quad \ldots \quad 30$$

calculated for 91 I compounds with two activating groups at A and B using equation 28,and the results (equation 30) were far from satisfactory. The correlation coefficient was good but the slope in equation 30 is far from unity. For lipophilic compounds the deviation between calculated and observed log P' values is as much as + 0.7 log units while that for hydrophilic compounds is as much as -3.7 log units. These deviations were ascribed to interaction between the two contiguous A and B groups in the I compounds. Accordingly equation 28 was expanded to 31 to include a $\Pi'_{interaction}$ term (reference 84 page 787) to compensate for interaction between the contiguous groups. Employing equation 31 log P' values were

$$\log P' = \Sigma\Pi' + \Sigma\Pi'_{interaction} - 1.30 \quad \ldots \quad 31$$

calculated for 325 I compounds with one and two activating groups. The relationship between log P' (calc) and log P' (obs) is expressed by equation 32.

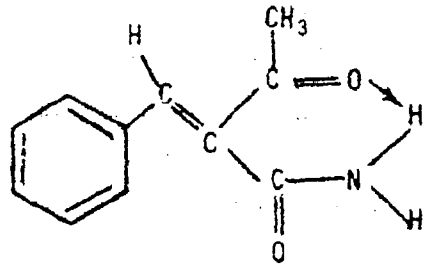
$$\log P' \text{ (calc)} = 0.98 \log P' \text{ (obs)} + 0.08$$
$$n = 325,\ r = 0.98 \quad \ldots \quad 32$$

Directing attention similarly to log P values for the I compounds in the system 1-octanol-water it again was found necessary to introduce a $\Pi_{interaction}$ term (ref. 84 page 795) into equation 27 to give equation 33.

$$\log P = \Sigma\pi + \Sigma\pi_{interaction} \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad 33$$

$$\log P\ (calc) = 1.07 \log P\ (obs) - 0.16$$
$$n = 103,\ r = 0.97 \quad \ldots\ldots\ldots\ldots \quad 34$$

Log P values were calculated for 103 compounds using equation 33 and the relationship between log P (calc) and log P (obs) is given by equation 34. Benzalacetoacetamide probably exists in the hydrogen bonded form VI.



VI

Calculation of log P for this compound by equation 33 (including $\pi_{interaction}$) gave a calculated value of log P which differed from the observed value (ref. 84 page 754) by -0.45 log units. (The deviation between log P' (calc) and log P' (obs) was -1.76 log units.)

Folding of the molecule by whatever means generally leads to a more hydrophilic compound than would be predicted from either equation 31 or 33. Hansch and Anderson (19) developed π constants in three different ways.

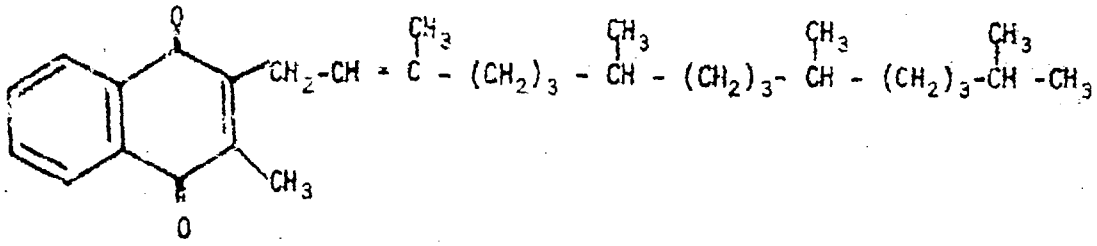
1) $\pi_X = \log P_{C_6H_5CH_2CH_2CH_2X} - \log P_{C_6H_5CH_2CH_2CH_3}$

2) $\pi_X = \log P_{C_6H_5CH_2X} - \log P_{C_6H_5CH_3}$

3) $\pi_X = \log P_{C_5H_{11}X} - \log P_{C_5H_{12}}$.

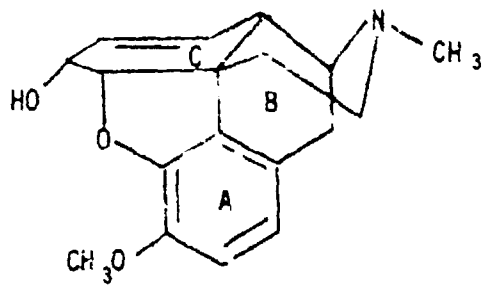
The π_X values developed by these three methods all differed, however there was a constant difference between the π_X values developed by methods 1 and 3. These workers ascribe this to a folding of the side chain in $C_6H_5CH_2CH_2CH_2X$

due to interaction of the side chain dipole moment with the π electrons of the aromatic ring. This, when reinforced by hydrophobic bonding, could lead to a folded molecule and this leads to greater hydrophilicity than would be expected. Currie *et al* (111) observed that log P' for vitamin K_1, VII, was much lower than would be predicted from log P' for 2,3-dimethyl-1, 4-naphthoquinone. This too may be due to folding of the long side chain due to hydrophobic bonding.

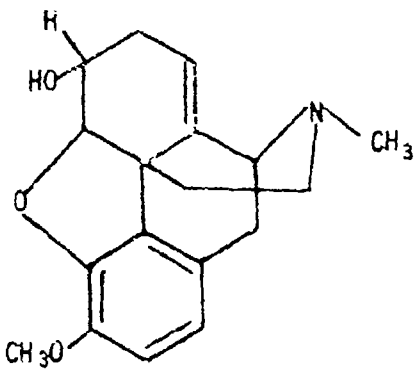


VII

The phenanthrylene oxide bridge and the ethanamine chain in codeine impart rigidity to the backbone (rings A and B) of codeine, VIII, and neopine, IX. The Δ^{7-8} double bond in codeine leads to a puckered molecule while the Δ^{8-14} double bond of neopine leads to a nearly planar structure (ref 84 pages 104-109). Log P''' (ether-water) for codeine is + 0.09 while that for neopine is - 0.43. Obviously the folding of the codeine molecule is blanketing part of the molecule from interaction with the solvent. Folding and hydrogen bonding may account for the anomalies



VIII



IX

present in the log P values for the hydrochlorides of the benzomorphans listed in Table 1.

Substituent Molar Refraction Constants, MR (92)

E_s has been successfully used in some equations for evaluating steric effects in the interaction of drugs with biological systems, however the steric requirements are often of the bulk type so Hansch, for want of a better parameter, has used group molar refractivities which are listed in reference 92. MR, besides evaluating bulk steric effects, is also directly related to polarizability so it should be used with caution. Another limitation to the use of MR in this work is that there is some collinearity between MR and π (58,67,99) and other variables (61,70). MR has been scaled down by a factor of 0.1 for use by Hansch in references 53,58,59,61,66,67,73 . This makes MR more equiscalar with π (53).

TABLE 1

LOG P VALUES FOR SALTS OF SOME BENZOMORPHAN DERIVATIVES IN 1-OCTANOL-WATER

COMPOUND	LOG P
2-Methyl-6,7-benzomorphan hydrochloride	-1.65
2-Methyl-5,9-diethyl-6,7-benzomorphan hydrochloride	-0.98
2-Hydroxy-2-methyl-5-ethyl-6,7-benzomorphan hydrochloride	-1.28
2-Hydroxy-2,9-dimethyl-5-ethyl-6,7-benzomorphan hydrochloride	-0.97
2-Hydroxy-2,5-dimethyl-9-ethyl-6,7-benzomorphan hydrochloride	-1.12

QSAR EQUATIONS DERIVED FROM EXTRATHERMODYNAMIC CONSIDERATIONS

Hansch (1,2,3), assuming the rates of metabolism and elimination to be zero or constant and steric factors to be insignificant, has developed, from extrathermodynamic (69) considerations, equation 35 relating biological response to hydrophobic and electronic factors. In equation 35 biological response is usually recorded as $\log \frac{1}{C}$ where C is the molar concentration of the drug eliciting the desired response. Under specific conditions

$$\log \frac{1}{C} = -k_1 \pi^2 + k_2\pi + k_3\sigma + k_4 \quad \ldots \quad 35$$

(see ref. 2 and 3) this equation may reduce to the simpler equations 36-39.

$$\log \frac{1}{C} = a\pi + b \quad \ldots \quad 36$$

$$\log \frac{1}{C} = -a\pi^2 + b\pi + c \quad \ldots \quad 37$$

$$\log \frac{1}{C} = \rho\sigma + c \quad \ldots \quad 38$$

$$\log \frac{1}{C} = a\pi + \rho\sigma + c \quad \ldots \quad 39$$

If steric factors are not insignificant then E_S or MR can be introduced into these equations while σ^+, σ^-, σ^* or E_R can be substituted for σ as the situation demands. The data for $\log \frac{1}{C}$,π,σ, E_S etc.can be submitted to multiple regression analysis to derive the best values for a, b, c etc. Deriving values for all these equations and examination of the correlation coefficients will indicate which equation best accounts for the biological activity. Table 2 summarizes the types of equations that have been developed so far while Table 3 presents some of the biological processes that have been examined by this method. This approach will be illustrated by examining ten type equations for QSAR.

Four basic steps must be followed in the development of the best equation to represent the biological process and they are:

1) Equations are developed by linear regression analysis relating the logarithms of the biological responses of a family of drugs to each of π, σ etc., E_S and MR. Even if the correlation coefficient and standard deviation are not good, the correlation coefficient, r, shows the relative importance of each of these

TABLE 2

SUMMARY OF EQUATIONS RELATING BIOLOGICAL RESPONSES TO VARIOUS COMBINATIONS OF SUBSTITUENT CONSTANTS

Equations Involving	References	Equations Involving	References
π (or log P)	7,13,18,23,24, 25,35,38,40, 42,45,46,64,68	Π and μ	62
		σ^* and E_S	15
σ	34,37	E_S and MR	66
σ^+	21,34,41	π^2, π and σ	35
σ^-	30,40,41	π^2, π and $\sigma\cdot$	21
		π^2, π and E_R	28
σ^*	15,61,68	π^2, π and $\mathcal{F}$	59
$\sigma\cdot$	21,28,41	(log P)^2, log P and E_S	68
E_S	15,22,23,37, 39,40,51,53	π^2, π and D	40
π^2 and π (or (log P)^2 and log P)	24,35,48, 50,57,68	(log P)^2, log P and pK_A	24
π and σ	14,15,17,22, 35,37	π, σ and E_S	15,27
		log P, σ^* and E_S	61
π and σ^+	34,67	π, $\mathcal{F}$ and MR	59
π and E_S	24,37,64	π, MR and D	69
π and MR	67		

TABLE 3

SOME BIOLOGICAL PROCESSES SUBJECTED TO QUANTITATIVE STRUCTURE-ACTIVITY ANALYSIS

Test Medium	Biological Action	Drugs	No. of Test Drugs	References
A				
Adenasine Deaminase (Bovine Gut Mucosa)	Inhibition of	9-(1-alkyl-2-hydroxyethyl) adenines	9	50
Aerobacter aerogenes	Inhibition of Growth	β-Nitrostyrenes	14	84
Antibody - Antigen Interaction	Inhibition of	Benzoate anions	22	26,53
A. oleracea	Kill	N-Alkylpyridinium Halides	7	35
		Quinones	10	35
Arbacia Eggs	Inhibition of Cell Division	Barbiturates	19	16
Aspergillus niger	Inhibition of Growth	Benzyl alcohols	19	35
		Isothiocyanates	13	35
		β-Nitrostyrenes	8	84
		Phenols	26	35
		α,β-Unsaturated Ketones	19	35
	Kill	Carboxylate ions	8	35
Aspergillus solani	Inhibition of Growth	Imidazolines	15	35

B

Bacillus subtilis	Inhibition of Growth	Rifamycin B amides and hydrazides	24 and 41	61
Bacteria	Inhibition of Luminescence	Alcohols and Urethanes	5 and 8	16
Bacteria and Fungi	Antimicrobial Activity	Esters of p-Hydroxybenzoic acid	4 and 26	42
Barnacle Larvae	Narcosis	Alcohols	14	44
Beef Erythrocyte Carbonic Anhydrase	Inhibition of	4-Subst-benzenesulfonamides	12	50
Beef Liver Mitochondria	Deamination of Amines	Benzylamines	7	24
Blood Clot	Fibrinolytic activity	Salicylic and Benzoic Acids	49	33
		Benzoates, Salicylates, anthranilates and cycloprcpanecarboxylate anions	9 and 22	64
Blood Plasma	Fibrinolytic activity	Benzoic Acids	15	33
		Salicylic Acids	13	33
Botrytis allii	Curling of hyphae	Griseofulvin Analogs	22	35
	Inhibition of Growth	β-Nitrostyrenes	20	84
Botrytis cinerea	Inhibition of Growth	β-Nitrostyrenes	8	84
Bovine hemoglobin	Binding of	Miscellaneous compounds	17	44
Bovine Muscle Succinate Oxidase	Inhibition of	Miscellaneous compounds	14	44

C				
Candida albicans	Inhibition of Growth	Carboxylate Ions	6	35
		Diamines	19	35
		Hydroxybenzoic Esters	7	35
		β-Nitrostyrenes	14	84
		Pyrimidines	8	35,44
TA3 Carcinoma Cells (Mice)	Antitumor Activity	Diaminopyrimidines	10	68,75
Carcinoma of Nasopharynx	Antitumor Activity	N-Acyltriamines	28	75
Cat	Analgesia	Morphine Alkaloids	14	84
	Antagonism to adrenaline	N-Substituted-2-bromoalkylamines	10	22
	Excitant Activity	Morphine Alkaloids	14	84
Cattle Liver Amine Oxidase	Deamination of Amines	Primary Amines	8	24
Cell Culture Nasopharyngeal Carcinoma	Antitumor Activity	Acyldiamines	28	68,75
Chloroplasts	Inhibition of Hill Reaction	Anilides	25	14
		Phenyldimethyl-ureas	12	14
		N-Phenylisopropyl-carbamates	9	14
Cholinesterase	Inhibition of	Diethyl Phenyl Phosphates	5	30
		Methylcarbamates, Diethylphenylphosphates, Alkylphosphonic acid esters and Phosphoramidates	8 and 30	15

Cholinesterase (Human Plasma)	Inhibition of	$R\overset{+}{N}(CH_3)_3$	7	50
Cholinesterase (Rabbit Plasma)	Inhibition of	$(CH_3)_3\overset{+}{N}(CH_2)_n\overset{+}{N}(CH_3)_3$	7	50
D				
Dihydrofolate Reductase	Inhibition of	Pyrimidines and Triazines	10 and 12	17
		Triazines	83 and 244	58,73
Dihydrofolate Reductase (*E.coli*)	Inhibition of	Triazines	12 and 15	17,40
Dihydrofolate Reductase (Pigeon Liver)	Inhibition of	Pyrimidines	12	40,50
Dihydrofolate Reductase (Tumor)	Inhibition of	Triazines	244	69
E				
Equine Liver Dehydrogenase	Inhibition of	Carboxyamides	6	50
Escherichia coli	Inhibition of Growth	Chloramphenicols	10	28
		β-Nitrostyrenes	23	84
F				
Flies	Synergistic action	1,3-Benzdiazoles with Carbaryl	16	21
Frog Flexor Reflex	Stimulation of	Benzaldehydes	13	84
		Catechol monomethyl ethers	13	84
		α-Haloacetophenones	13	84
		β-Nitrostyrenes	29	84
		Tetrahydro-1,4-naphthoquinones	12	84

Frog Flexor Reflex	Stimulation of	α,β-Unsaturated compounds	46	84
Frog Sciatic Nerve	Inhibition of	Alcohols	8	44
Fusarium bulbigenum	Fungitoxicity	β-Nitrostyrenes	8	84
G				
G. cingulata	Inhibition of Growth	Imidazolines	15	35
Goldfish	Narcosis	Alcohols	8	44
Guanine Deaminase	Inhibition of	9-Phenylguanines	32 and 96	51,73
Guinea Pig Complement	Inhibition of	Benzamidines	108	53,73
		Benzylpyridinium ions	69	67,73
Guinea Pig Eye	Blepharospasm	α,β-Unsaturated compounds	10	84
Guinea Pig Ileum	Cholinergic ED_{50}	Choline Esters	6	44
	In vitro Inhibition	Diphenylhydramines	30	27
Guinea Pigs	Anaesthetic action	2-Diethylaminoethyl benzoates	8	3
	In vivo Histamine Response	Diphenylhydramines	22	27
H				
Hansenula anomala	Inhibition of Growth	Phenyl Methacrylates	10	35
Horse Heart Cytochrome	Denaturation of	Alcohols	5	44
House flies	Toxicity of	Diethyl phenyl phosphates	14	3
Human Liver Mitochondria	Monamine Oxidase Inhibition	N-(phenoxyethyl) cyclopropylamines	9	27

I				
Influenza B. Virus	Inhibition of Multiplication	Benzimidazoles	15	44
L				
Letinus lepideus	Inhibition of Growth	Hydroxy Compounds	5	35
		Ammonium Ions	6	35
Lewis Lung Carcinoma	Antitumor activity	Nitrosoureas	8 and 13	57,68
L 1210 Leukemia (Mice)	Antileukemia Activity	Nitrosoureas	22	68
		Triazines	10	68
Lobster Axon	Resting Potential Change	Alcohols	5	44
M				
Macrosporum sarcinaeforme	Inhibition of Growth	Imidazolines	15	35
M. aureus	Inhibition of Growth	Rifamycin B amides and hydrazides	24 and 41	61
M. fructicola	Inhibition of Growth	N-Alkylethylene-thioureas	5	35
" " (Spores of)	Lethal Dose	Benzoquinones	10	44
M. tuberculosis	Inhibition of Growth	Phenols	14	44
Mice	Anaesthetic Activity	Aliphatic Ethers	26	46a
		Gaseous anaesthetics	30	62
	Analgesia	Morphine Alkaloids	9	84
	Antagonism of Adrenaline	N-Substitut 2-bromoalkylamines	9	22
	Antileukemic Activity	Nitrosoureas and Imidazole-carbosyamides	10 and 22	41

Mice	Convulsant Action	Morphine Derivatives	14	84
	Lethal Dose	Morphine Derivatives	6 and 14	84
	Hypnosis of	Acetylenic Alcohols	8	44
		Alkyl-aryl-ureas	23	44
Mice, Rabbits and Rats	Hypnosis of	Barbiturates	102	18
		Non-barbiturates	73	18
		Thiobarbiturates	25	18
Monamine Oxidase (Beef Liver Mitochondria)	Inhibition of	9-Alkyl-β-carbolines	8	50
		Isonicotinic acid hydrazides	6	50
" " (Human Liver Mitochondria)	Inhibition of	Alcohols	5	50
" " (Mouse Liver Mitochondria)	Inhibition of	Isonicotinic Acid Hydrazides	6	50
" " (Rat Brain)	Inhibition of	$HOCH_2CH(NH_2)CONHNHR$	7	50
Monilia fructicola	Kill	Quinones	10	35
Mosquito larvae	Toxicity to	Benzoic Acids	14	3
Mouse Eye	Blepharospasm	α,β-Unsaturated Compounds	13	84
Mouse TA 3 Carcinoma	Antitumor Activity	Diaminopyrimidines	9	75
$\underline{N}$ NADH Oxidation	Inhibition of	Barbiturates	6	16

P

Penicillium cyclopium	Inhibition of Growth	Isothiocyanates	6	35
Penicillium Lumber Mould	Inhibition of Growth	β-Nitrostyrenes	8	84
Pepsin (Porcine)	Inhibition of	Alcohols	8	50
Phytophthora infestans	Kill	N-Alkylpyridinium Halides	8	35
Plants	Elongation of	Phenoxyacetic acids	20 and 35	1,2
Pseudomonas aeruginosa	Inhibition of Growth	β-Nitrostyrenes	17 and 19	84

R

Rabbit Kidney	Inhibition of Oxidation of Indophenol	Alcohols and Ketones	8	44
Rabbit Kidney Reductase	Reduction of Acetophenones	4-Substituted Acetophenones	10	34
Rabbit Liver Amine Oxidase	Deamination of amines	Primary and Secondary amines	10 and 30	24
Rabbit Liver	Inhibition of Guanine Deaminase	9-(X-Phenyl)guanine	32	51
Rabbits	Glucuronamide formation	Aliphatic Alcohols	11	24
	Hippuric Acid formation	Benzoic Acids	8	24
	Progestational Activity	Δ^6-6-Substituted-Progesterones	13	59
	Respiratory Depression	Morphine Derivatives	9	84
Rat Brain	Inhibition of Oxygen Consumption	Barbiturates	10	16

Rat Liver	Carcinogenic activity of	Dimethylaminoazo-benzenes	41	3
	Monamine Oxidase Inhibition	N-(phenoxyethyl) cyclopropylamines	18	27
Rat Microsomes	Demethylation of	Tertiary amines	18	10
Rats	Antagonism of Adrenaline	N-Substituted-2-bromoalkyl amines	11	22
	Antagonism of Adrenaline & Noradrenaline	N,N-Dimethyl-2-bromophenethylamines	22	22
	General Depression	Morphine Derivatives	14	84
Rat Tumor	Inhibition of Guanine Deamin-ase	9-(X-Phenyl)guanine	18	51
Red Blood Cells	Hemolysis of	Aliphatic Acids	9	44
Rodents	Thyroxine-like Activity	Thyroxin analogues	9	3
S				
Sheep Liver Succinate Oxidase	Inhibition of	Miscellaneous Compounds	14	44
Staphylococcus albus	Inhibition of Growth	β-Nitrostyrenes	15	84
		α,β-Unsaturated Compounds	26	84
Staphylococcus aureus	Inhibition of Growth	Alcohols	9	44
		Benzylammonium Chlorides	45	44
		Chloromycetin analogs	11	2
		β-Nitrostyrenes	35	84
		Penicillins	8	9
		Phenoxypenicillins	21	44

Staphylococcus aureus	Inhibition of Growth	α,β-Unsaturated Compounds	35 and 52	84
Staphylococcus typhosa	Inhibition of Growth	Aromatic Amines	15	44
		Phenols	35	3
Streptococcus faecalis	Inhibition of Growth	Rifamycin B amides and hydrazides	24 and 41	61
Streptococcus hemolyticus	Inhibition of Growth	Rifamycin B amides and hydrazides	24 and 41	61
T				
Tadpoles	Narcosis of	Alcohols	8	44
Trichomonas vaginalis	Inhibition of Growth	β-Nitrostyrenes	14	84
Trichophyton asteroides	Inhibition of Growth	β-Nitrostyrenes	8	84
Trichophyton gypseum	Inhibition of Growth	Phenols	11	35
Trichophyton interdigitale	Inhibition of Growth	Alkylpyrazoles	6	35
		Carboxylate Ions	14	35
		Diamines	22	35
Tongue	Relative Sweetness	m-Nitroanilines	9	34
Trypsin (Beef Pancreas)	Inhibition of	$H_2N(CH_2)_nNH_2$	6	50
Tyrosine Hydroxylase (Beef Adrenal Medullae)	Inhibition of	Carboxyamides	6	50
U				
Urease (Sword Bean)	Inhibition of	R-CONHOH	11	50

V				
Valyl-t-RNA Synthetase (*E.coli*)	Inhibition of	RNH_3^+ Salts	10	50
Venturia inaequalis	Fungitoxicity	β-Nitrostyrenes	8	84
	Kill	N-Alkylpyridinium Halides	8	35
W				
Walker 256 Carcinoma	Antitumor Activity	Aromatic Nitrogen Mustards	9	75
		Aziridinyl-l-phosphinyl carbamates	10	68,75
		Aziridines	8	68
		Bis-(l-aziridinyl)-phosphinyl carbamates	10	75
Whale Myoglobin	Denaturation of	Alcohols	7	44
X				
Xanthine Oxidase	Inhibition of	9-(X-Phenyl) guanines	30	66,73

single parameters to the biological response.

2) Equations are next developed by multiple regression analysis relating the logarithms of the biological responses to combinations of two of the parameters then three etc. and the correlation coefficients examined to determine the one with the highest value and the lowest standard deviation.

3) In certain cases a dummy, D, or indicator, I, parameter may prove helpful.

4) It is difficult to draw any conclusions from the above data if collinearity exists between two or more of the parameters in an equation (70) so the various parameters must be plotted against each other to determine if collinearity exists between any of them.

Ten examples are now presented to illustrate this method.

1) Equation Involving only π or Log P (46)

Most of the equations so far developed, in which only π or log P are involved, are related to the binding of organic compounds to proteins (13,20) and enzymes (20,191) and the growth inhibition of bacteria (46). The binding of compounds to both bovine serum albumin (6) and bovine hemoglobin (13) is dependent solely upon log P; electronic and steric factors having no influence. A family of 19 phenols was used in examining the binding, $\frac{1}{C}$ (C = the molar concentration to produce a 1:1 complex), to bovine serum albumin (equation 40), while seventeen phenols, amines and neutral compounds were used in the analogous study with bovine hemoglobin (equation 41).

$$\log \frac{1}{C} = 0.681 \log P + 2.489$$
$$n = 19,\ r = 0.962,\ s = 0.133 \quad \ldots\ldots\ldots \quad 40$$

$$\log \frac{1}{C} = 0.713 \log P + 1.512$$
$$n = 17,\ r = 0.950,\ s = 0.160 \quad \ldots\ldots\ldots \quad 41$$

2) Equation Involving only σ, σ^+ or σ^*

Few examples have been encountered where some biological process is related solely to σ or $\sigma\cdot$ etc. The reason for involvement of only this parameter appears to be due to steric interaction between the substituents of the

compound with the enzyme or lipoprotein membranes (37). One example involving only σ is in the enzymatic hydrolysis of phenyl sulphates (equation 42).

$$\log \frac{1}{K_m} = 0.930\,\sigma + 2.522$$
$$n = 10,\ r = 0.931 \quad \ldots\ldots\ldots\ldots\ldots \quad 42$$

McMahon *et al* (118) studied the enzymic reduction of a series of ten 4-substituted-acetophenones which can involve resonance between the substituent and the ketone group (see page 15). Hansch developed equations 43 and 44 from the data of McMahon *et al* for the rate of reduction, k_0, by rabbit kidney reductase (34).

$$\log k_0 = 2.042\,\sigma + 1.173$$
$$n = 10,\ r = 0.862,\ s = 0.467 \quad \ldots\ldots\ldots \quad 43$$

$$\log k_0 = 1.514\,\sigma^+ + 1.480$$
$$n = 10,\ r = 0.914,\ s = 0.390 \quad \ldots\ldots\ldots \quad 44$$

Obviously equation 44 (involving σ^+) fits the data better than does equation 43 (involving σ).

In developing equations for the bacterial inhibition of a number of bacteria by thirty-six rifamycin B amides, Hansch (61) found that equation 45 for *Staphylococcus hemolyticus* involved only σ^*.

$$\log \frac{1}{C} = -0.93\,\sigma^* + 7.83$$
$$n = 36,\ r = 0.858,\ s = 0.276 \quad \ldots\ldots\ldots \quad 45$$

3) Equations Involving only E_S

The superb work of Pauling and his collaborators (119) in the hapten-antibody interaction has been analyzed by Kutter and Hansch (26). The 18 haptens are all simple substituted benzoate ions. Neither π nor σ appears to exert any influence, however, the steric effects, E_S^o, E_S^m and E_S^p of the substituents in the ortho, meta and para positions do as is manifest in equation 46.

$$\log K_{rel} = 1.199\,E_S^o + 0.368\,E_S^m - 0.467\,E_S^p - 1.406$$
$$n = 18,\ r = 0.981,\ s = 0.217 \quad \ldots\ldots\ldots \quad 46$$

4) Equations Involving $(\text{Log P})^2$ and Log P (48)

In contrast to the inhibition of enzymes which is related only to Log P or π, the equations for biological responses of drugs upon living organisms usually involve $a(\log P)^2$ or π^2 term (see equations 35 and 37). Experience has shown that highly hydrophilic compounds in a series of drugs generally have little if any biological activity and as the lipophilicity increases so does the response. Similarly highly lipophilic compounds in the same family of drugs show little if any biological activity and as the hydrophilicity increases so does the biological activity. The activity from either end of the scale does not increase indefinitely but reaches a maximum. Somewhere between P = 0 and P = ∞ for a given family of drugs acting in a given biological system will be an ideal P called P_0 where the biological activity will be a maximum. A greater number of molecules of the drug with P_0 would reach the site of action in the test interval of time than would drugs having another P value. The shape of the above plot of log BR against log P leads to a parabola: hence the $(\log P)^2$ or π^2 in equations 35 and 37.

Hansch, Smith, Engle and Wood (43,68) studied the antileukemic activity of a family of nitrosoureas against BDF_1 mice inoculated intracerebrally with 10^4 cells of L 1210 leukemia which led to equations 47 and 48. Unfortunately no σ or E_S values were available for some of the

$$\log \frac{1}{C} = -0.282 \log P + 4.639$$
$$n = 22,\ r = 0.886,\ s = 0.189 \quad \ldots\ldots\ldots \quad 47$$

$$\log \frac{1}{C} = -0.0568 (\log P)^2 - 0.0689 \log P + 4.527$$
$$n = 22,\ r = 0.922,\ s = 0.163 \quad \ldots\ldots\ldots \quad 48$$

substituents in the nitrosoureas so linear equations involving only σ and E_S could not be developed. However the high correlation coefficient of equation 47 indicates that the biological activity of the nitrosoureas are highly dependent upon log P. The increase in the correlation coefficient upon introduction of a $(\log P)^2$ term is statistically significant. The negative coefficient of log P in equation 47 is very interesting. The more hydrophilic the compound becomes, the greater is its activity. This, however, is only half the problem. Toxicity must also be kept to a minimum.

LD_{10} values for the nitrosoureas on mice were used as approximate values of toxicity and equations 49 and 50 were developed by the method of least squares. Log P_0 (see page 83) for toxicity is 1 log unit higher than that for potency of the nitrosoureas. This indicates that one should

$$\log \frac{1}{C} = 0.210 \log P + 4.232$$
$$n = 28,\ r = 0.737,\ s = 0.232 \quad \ldots\ldots \quad 49$$

$$\log \frac{1}{C} = -0.0688 (\log P)^2 + 0.0593 \log P + 4.066$$
$$n = 28,\ r = 0.809,\ s = 0.206 \quad \ldots\ldots \quad 50$$

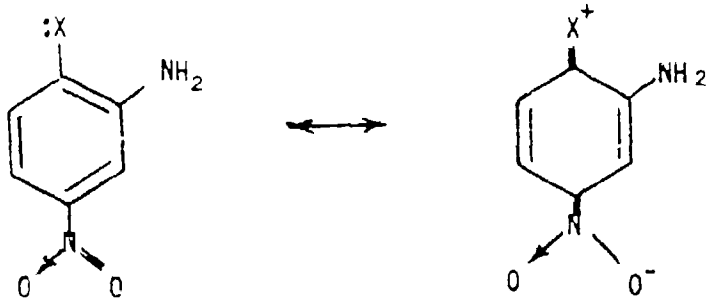
be able to make more potent nitrosoureas by reducing their lipophilicity and at the same time lowering toxicity.

5) Equations Involving π and σ etc.

Miller and Hansch (17) have applied regression analysis to the data of Baker and Shapiro (120) on the inhibition of dihydrofolate reductase by pyrimidines. Equation 51 summarizes this relationship.

$$\log \frac{1}{C} = +0.457\ \pi\ -5.820\ \sigma\ -6.951$$
$$n = 16,\ r = 0.903,\ s = 0.741 \quad \ldots\ldots \quad 51$$

Turning next to the work of Blanksma and Hoegen (121) on the sweetness of some substituted-m-nitranilines, Xa ⟷ Xb.



Xa Xb

Hansch (34) developed equations 52 and 53 relating the relative sweetness (RS) to hydrophobic (π) and electronic (σ or σ^+) constants. Considering

$$\log RS = 1.610\,\pi - 1.831\,\sigma + 1.729$$
$$n = 9,\ r = 0.936,\ s = 0.282 \quad \ldots\ldots\ldots\ldots \quad 52$$

$$\log RS = 1.434\,\pi - 1.026\,\sigma^+ + 1.584$$
$$n = 9,\ r = 0.972,\ s = 0.190 \quad \ldots\ldots\ldots\ldots \quad 53$$

the possibility of resonance between the X and the NO_2 groups in Xa $\leftrightarrow$ Xb it is not surprising to find that equation 53 involving σ^+ (see page 15) gives the higher correlation coefficient and lower standard deviation.

6) Equations Involving π and E_S

Two cases will be examined under this heading:

1) the fibrinolytic activities of substituted benzoate ions (64)
2) the rate of metabolic change, MR, of m- and p-substituted benzylamines in the presence of beef liver mitochondria (24).

The fibrinolytic activity of the 3- and 4-substituted benzoic acids is linear with respect to log P but when ortho-substituted derivatives are included a steric factor E_S is required. This led to equation 54 for the ortho-substituted-benzoic acids.

$$\log \frac{1}{C} = 0.48 \log P + 0.44\,E_S + 1.36$$
$$n = 16,\ r = 0.885,\ s = 0.210 \quad \ldots\ldots\ldots\ldots \quad 54$$

In the second case cited (24) the metabolic change (MR) in beef liver mitochondria when treated with meta-substituted benzylamines is linear with respect to log P (equation 55) but that of the para-substituted benzylamines is not (equation 56). Insertion of an E_S term into the equation (eq. 57) for the para-substituted benzylamines markedly improved the correlation coefficient. Combining the m- and p-substituted benzylamines gave equation 58.

$$\log MR = 0.452 \log P + 1.767$$
$$n = 7,\ r = 0.954,\ s = 0.085 \quad \ldots\ldots\ldots \quad 55$$

$$\log MR = 0.256 \log P + 1.229$$
$$n = 7,\ r = 0.278,\ s = 0.455 \quad \ldots\ldots\ldots \quad 56$$

$$\log MR = 0.753 \log P + 0.535\ E_S + 0.304$$
$$n = 7,\ r = 0.787,\ s = 0.327 \quad \ldots\ldots\ldots \quad 57$$

$$\log MR = 0.623 \log P + 0.683\ E_S + 0.554$$
$$n = 13,\ r = 0.874,\ s = 0.293 \quad \ldots\ldots\ldots \quad 58$$

7) Equations Involving π^2, π and σ etc.

Three biological processes will be examined under this heading.

1) The antifungal activities of phenols against *Aspergillus niger* (35).
2) The synergistic activity of some methylenedioxy compounds for the insecticide carbaryl (21).
3) The antibacterial activities of some chloramphenicols against gram-negative bacteria (28).

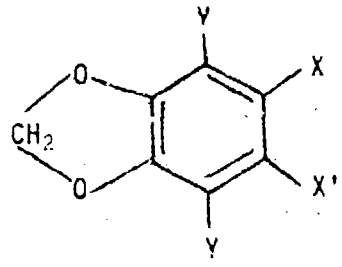
A variety of antifungal agents was examined by Hansch and Lien (35) against a large number of species of fungi and equations developed by the usual method. Of these, 12 had equations of the form of equation 36, seven gave equations of the form of equation 39 and thirty-two had equations of the form of equation 37. Equations for the action of phenols upon *Aspergillus niger*, phenyl methacrylates upon *Hansenula anomala* and $RR'NCSS^-Na^+$ upon *Botrytis cinerea* conformed to the general equation 35. The equations for the antifungal action against these three organisms are respectively equations 59-61.

$$\log \frac{1}{C} = -0.190\ (\log P)^2 + 1.859 \log P + 0.627\ \sigma - 0.092$$
$$n = 18,\ r = 0.975,\ s = 0.160 \quad \ldots\ldots \quad 59$$

$$\log \frac{1}{C} = -0.120\ (\log P)^2 + 1.234 \log P - 0.880\ \sigma + 0.878$$
$$n = 10,\ r = 0.958,\ s = 0.069 \quad \ldots\ldots \quad 60$$

$$\log \frac{1}{C} = -0.282\ (\log P)^2 \quad -0.207 \log P - 1.531\ \sigma + 5.063$$
$$n = 9,\ r = 0.921,\ s = 0.278 \quad \ldots \quad 61$$

It is well known that many methylenedioxy compounds are synergists for the insecticide carbaryl (1-naphthyl-N-methylcarbamate). The methylenedioxy compounds, XI, with the greatest synergistic activity were those containing nitro and methoxy groups. In nucleophilic and



XI

electrophilic substitutions these functions have opposite effects. However, in certain homolytic (free radical) substitutions, nitro and methoxy are amongst the strongest promoters of reaction (116 page 57). This prompted the use of $\sigma\cdot$ in the equations. Equations 62-65 were developed by Hansch (21) from the synergistic activities, SR5*, reported by Wilkinson (122). Hennessy (123) from the old classical approach to structure-activity postulated that the synergistic mechanism involved the loss of one hydrogen

$$\log SR5 = 0.940\sigma\cdot + 1.963$$
$$n = 13,\ r = 0.603,\ s = 0.334 \quad \ldots \quad 62$$

$$\log SR5 = 0.070\ \pi + 0.917\ \sigma\cdot + 2.050$$
$$n = 13,\ r = 0.638,\ s = 0.338 \quad \ldots \quad 63$$

---

* SR5 is the synergistic activity when the ratio of synergist to insecticide is 5:1.

$$\log SR5 = -0.115\ \pi^2 + 0.348\ \pi + 2.146$$
$$n = 13,\ r = 0.500,\ s = 0.380 \quad \ldots\ldots \quad 64$$

$$\log SR5 = -0.195\ \pi^2 + 0.670\ \pi + 1.316\ \sigma\cdot + 1.612$$
$$n = 13,\ r = 0.929,\ s = 0.171 \quad \ldots\ldots \quad 65$$

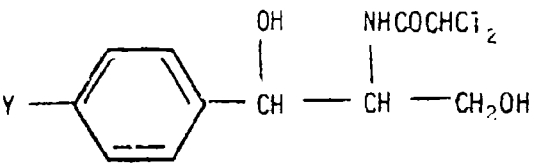
from the methylenedioxy group as a hydride ion. This prompted Hansch (21) to develop equations 65 and 67 involving σ^+ and σ_I in place of $\sigma\cdot$.

$$\log SR5 = -0.128\ \pi^2 + 0.032\ \pi + 0.945\ \sigma_I + 1.851$$
$$n = 13,\ r = 0.852,\ s = 0.242 \quad \ldots\ldots \quad 66$$

$$\log SR5 = -0.113\ \pi^2 + 0.374\ \pi - 0.166\ \sigma^+ + 2.184$$
$$n = 13,\ r = 0.532,\ s = 0.392 \quad \ldots\ldots \quad 67$$

From a consideration of correlation coefficients and standard deviations equation 65 with its free radical mechanism would appear to fit the facts better than one involving the loss of a hydride ion.

Homolytic constants, E_R, have been developed by Yamamoto and Otsu (110) (see page 15) and these were used by Hansch, Kutter and Leo (28) in the development of equations from the data of Garrett *et al* (124) on the inhibition of *E. coli* by a series of amphenicols, XII. The exploratory equation 68, developed by Hansch *et al* (2),



XII

demonstrated that activity, A, depended upon both electronic and hydrophobic properties. The low correlation for equation 68 prompted

Hansch *et al* (28) to employ the homolytic substituent constant, E_R in equations 69-73. Examination of the correlation coefficients for the

$$\log A = -0.74\ \pi^2 + 0.36\ \pi + 1.82\ \sigma_m + 0.62$$
$$n = 10,\ r = 0.824,\ s = 0.555 \quad \ldots . \quad 68$$

$$\log A = 2.744\ E_R + 0.931$$
$$n = 8,\ r = 0.820,\ s = 0.243 \quad \ldots . \quad 69$$

$$\log A = 0.145\ \pi + 1.289$$
$$n = 8,\ r = 0.317,\ s = 0.403 \quad \ldots . \quad 70$$

$$\log A = 0.227\ \pi + 3.069\ E_R = 0.769$$
$$n = 8,\ r = 0.954,\ s = 0.140 \quad \ldots . \quad 71$$

$$\log A = 0.187\ \pi + 3.419\ E_R - 0.235\ \sigma + 0.786$$
$$n = 8,\ r = 0.970,\ s = 0.127 \quad \ldots . \quad 72$$

$$\log A = -0.053\ \pi^2 + 0.231\ \pi + 2.865\ E_R + 0.846$$
$$n = 8,\ r = 0.957,\ s = 0.151 \quad \ldots . \quad 73$$

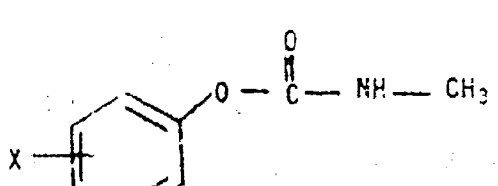
two single parameter equations 69 and 70 reveals that the electronic parameter is much more important than the hydrophobic one. While neither correlation is good the linear combination of these two parameters gives an excellent correlation. The addition of a π^2 term does not improve the correlation coefficient significantly for the limited range of compounds examined.

8) Equations Involving π, σ and E_S (15,27)

Two biological processes will be examined under this heading.

1) Inhibition of cholinesterase by carbamates and phosphate ester amides (15).
2) Antihistamines and monamine oxidase inhibitors (27).

Hansch and Deutsch (15) developed equations 74-83 from the data of Metcalf and Fukuto (125) for the inhibition of cholinesterase by some ortho-, meta- and para-substituted-carbamates, XIII. From equations 74-79 a number of salient features become apparent. There is a markedly different



XIII

4-substituted carbamates

$$\log \frac{1}{C} = 0.742\ \pi + 3.525$$
$$n = 22,\ r = 0.768,\ s = 0.458 \quad . \ . \ . \ . \ . \quad 74$$

$$\log \frac{1}{C} = -1.302\ \sigma + 4.202$$
$$n = 23,\ r = 0.404,\ s = 0.654 \quad . \ . \ . \ . \ . \quad 75$$

$$\log \frac{1}{C} = 0.714\ \pi - 0.868\ \sigma + 3.486$$
$$n = 23,\ r = 0.839,\ s = 0.399 \quad . \ . \ . \ . \ . \quad 76$$

3-substituted carbamates

$$\log \frac{1}{C} = 0.876\ \pi + 4.347$$
$$n = 30,\ r = 0.773,\ s = 0.592 \quad . \ . \ . \ . \ . \quad 77$$

$$\log \frac{1}{C} = -2.052\ \sigma + 5.673$$
$$n = 30,\ r = 0.511,\ s = 0.802 \quad . \ . \ . \ . \ . \quad 78$$

$$\log \frac{1}{C} = 0.784\ \pi - 1.405\ \sigma + 4.618$$
$$n = 30,\ r = 0.845,\ s = 0.508 \quad . \ . \ . \ . \ . \quad 79$$

2-substituted carbamates

$$\log \frac{1}{C} = 0.815\ E_S + 4.557$$
$$n = 7,\ r = 0.349,\ s = 1.306 \quad . \ . \ . \ . \ . \quad 80$$

$$\log \frac{1}{C} = 1.659\ \pi + 4.074$$
$$n = 7,\ r = 0.566,\ s = 1.040 \quad . \ . \ . \ . \ . \quad 81$$

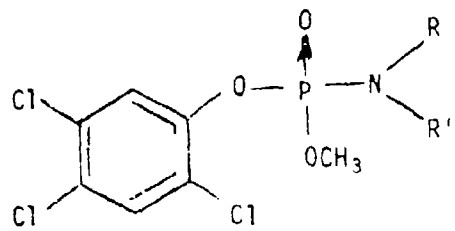
$$\log \frac{1}{C} = -1.393\ \sigma + 4.887$$
$$n = 7,\ r = 0.361,\ s = 1.300 \quad . \ . \ . \ . \ . \quad 82$$

$$\log \frac{1}{C} = 2.799\ \pi + 4.246\ \sigma + 3.845\ E_S + 2.542$$
$$n = 7,\ r = 0.962,\ s = 0.494 \quad \ldots . \quad 83$$

character in the activity of the meta and para isomers. If hypothetical meta and para isomers with $\pi = 0$ are considered the difference in activity of the two is 1.132 log units which is a difference of over 13 fold. Equations 74 and 75 as well as 77 and 78 reveal that π accounts for a much greater amount of the variance in $\log \frac{1}{C}$ than does σ. Moreover the positive coefficients associated with π in equations 74,76,77 and 79 indicate that the more lipophilic the compound the greater is the activity relative to that of the parent compound. The negative coefficients associated with σ in equations 75,76,78 and 79 indicate that electron donating groups on the phenyl group increase activity relative to the parent compound.

As might be expected steric factors become dominant in the 2-substituted-carbamates and the negative coefficient of E_S" in equation 83 reveals that the smaller the substituent the more active is the compound.

Hansch and Deutsch (15) also developed equations 84 to 89, based upon the data of Fukuto *et al*, for the inhibition, K, of cholinesterase by a homologous series of alkyl-2,4,5-trichlorophenyl-N-alkylphosphoramides, XIV.



XIV

---

* The larger the steric effect of the substituent the larger E_S becomes in a negative sense (66,68,27,44).

$$\log K = 2.709\ \sigma^* + 4.490$$
$$n = 8,\ r = 0.712,\ s = 0.816 \quad \ldots\ldots \quad 84$$

$$\log K = -1.011\ \pi + 6.567$$
$$n = 8,\ r = 0.608,\ s = 0.922 \quad \ldots\ldots \quad 85$$

$$\log K = 1.119\ E_S + 4.541$$
$$n = 8,\ r = 0.875,\ s = 0.563 \quad \ldots\ldots \quad 86$$

$$\log K = -3.913\ \sigma^* + 2.359\ E_S + 4.948$$
$$n = 8,\ r = 0.939,\ s = 0.438 \quad \ldots\ldots \quad 87$$

$$\log K = 0.001\ \pi + 1.119\ E_S + 4.540$$
$$n = 8,\ r = 0.875,\ s = 0.617 \quad \ldots\ldots \quad 88$$

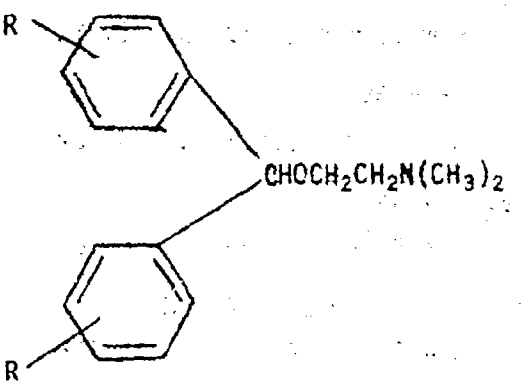
$$\log K = 0.007\ \pi - 3.913\ \sigma^* + 2.363\ E_S + 4.937$$
$$n = 8,\ r = 0.939,\ s = 0.490 \quad \ldots\ldots \quad 89$$

Of the single parameter equations, equation 86 gives the best fit thus demonstrating the importance of steric effects. Comparing equation 88 with 86 reveals that π contributes little if anything to the goodness of fit whereas equations 87 and 84 demonstrate that the reactivity parameter, σ^*, plays a significant role. The negative coefficients of σ^* in equations 87 and 89 reveal that electron donating R and R' groups promote activity.

Fuller *et al* (126) determined the inhibition, IG_{50}, of two types of monamine oxidases by N-(phenoxyethyl)cyclopropylamines. There is a marked difference in the activity of isomeric compounds when the substituent is in the meta or in the para position. Fuller attributed this to steric factors. Rutter and Hansch (27) evaluated the steric factors by E_S and developed equation 90 for this inhibitory process.

$$\log IG_{50} = 0.198\ \pi + 1.640\ \sigma + 0.702\ E_S + 4.153$$
$$n = 18,\ r = 0.945,\ s = 0.330 \ldots \quad 90$$

Rutter and Hansch (27) developed a similar series of equations for the antihistamine activity of a series of aryl-substituted-diphenylhydramines, XV. Harms and Nauta's (127) work was an *in vitro* study

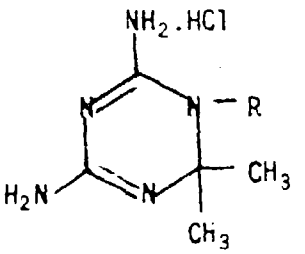


XV

while that of Ensor *et al* (128) was an *in vivo* study on guinea pigs. The equations for the two were quite independent of π and σ and involved $(E_S)^2$ and E_S terms. The equations derived from *in vitro* and *in vivo* data were surprisingly similar, so much so that both sets of data were encompassed by a single equation.

9) Equations Involving π^2, π and D (17, 40)

Miller and Hansch (17,40) applied substituent constants to Baker's results on the inhibition of dihydrofolate reductase by a homologous series of 1,3,5-triazines, XVI. In a number of cases electronic and steric factors



XVI

were not available for the large R substituents in XVI so a dummy parameter was introduced into the equation. The dummy parameter, D, is assigned the arbitrary value of 1.00 when a phenyl group is attached directly to the triazine ring and a value of 0.00 when an alkyl group is attached directly

to the triazine ring. This limited the study to two variables, π and D. Equations 91 to 93 summarize the results. Incorporation of D into

$$\log \frac{1}{C} = 0.74\ \pi + 4.25$$
$$n = 15,\ r = 0.816,\ s = 0.791 \quad \ldots\ldots\ldots \quad 91$$

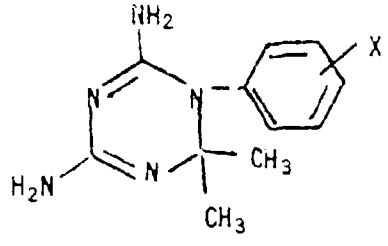
$$\log \frac{1}{C} = -0.23\ \pi^2 + 2.12\ \pi + 2.72$$
$$n = 15,\ r = 0.887,\ s = 0.657 \quad \ldots\ldots\ldots \quad 92$$

$$\log \frac{1}{C} = -0.28\ \pi^2 + 2.21\ \pi + 0.84\ D + 2.58$$
$$n = 15,\ r = 0.928,\ s = 0.553 \quad \ldots\ldots\ldots \quad 93$$

the equation gives a satisfactory correlation coefficient but the standard deviation is high. This is probably due to the inability of D to exactly evaluate electronic and steric factors. The positive coefficient of D indicates that a phenyl group attached directly to the triazine ring leads to greater inhibitory activity than does an alkyl group of equal lipophilicity.

10) Equations Involving π, MR and I (69)

As Baker's work expanded to include the inhibitory activities of more triazines upon dihydrofolate reductase from Walker 256 and L 1210 leukemia tumors, molecular refractions, MR, for substituent groups were being determined and Hansch and Silipo (58) developed equation 94 for eighty-three XVII compounds.



XVII

$$\log \frac{1}{C} = -0.13\ \pi_3^2 + 0.89\ \pi_3 + 0.15\ MR_4 + 6.62$$
$$n = 83,\ r = 0.905,\ s = 0.328 \quad \ldots\ldots \quad 94$$

The subscripts in equation 94 indicate the position of the substituent, X, in the phenyl group of structure XVII. At the termination of Baker's classical work, Silipo and Hansch (69) developed, in a stepwise fashion, a series of ten equations involving π_3 and MR_4 and six indicator parameters, I. For example I-1 made possible the merging into one equation of the activities of the XVII compounds against dihydrofolate reductase from the two types of tumors. The indicator value of I-1 was set at 1 for Walker enzyme data and at 0 for L 1210 enzyme. The complete equation for 244 compounds with structure XVII is 95.

$$\log \frac{1}{C} = -0.118\ \pi_3^2 + 0.680\ \pi_3 - 0.0243\ MR_4^2$$
$$+0.230\ MR_4 + 0.238\ I\text{-}1 - 2.530\ I\text{-}2$$
$$-1.991\ I\text{-}3 + 0.877\ I\text{-}4 + 0.686\ I\text{-}5$$
$$+0.704\ I\text{-}6 + 6.489$$
$$n = 244,\ r = 0.923,\ s = 0.377 \quad \ldots\ldots\ldots\ldots \quad 95$$

Summary

The versatility of this approach to the study of structure-activity relationships for bio-medical processes is impressive and has led to a dramatic step forward in the understanding of many of these processes, however it has some limitations. When log P values are calculated rather than measured compounds with intramolecular hydrogen bonding should be avoided. Also equations involving the electronic substituent constants σ, σ^+, σ^- and σ^* limit the number of families of compounds that can be included in the analysis to one. Equations 35 to 39 were developed on the premise that wastage of the drug by metabolism or elimination was not a significant factor.

While no attempt has been made to assess hydrogen bonding or loss of drug by metabolism and elimination, it is possible to develop equations separately for the rate of metabolism and of elimination as well as for the desired biological potency and the effect of log P or π on these three factors. Hansch *et al* (18) has done just this in the study

of the hypnotic effects of a series of barbiturates. Data were taken from the literature for the hypnotic effect of barbiturates in a variety of test animals and under different experimental conditions, however, the results were, in most cases, surprisingly similar. The activity against mice bears a parabolic relationship to log P as shown in equation 96. From data reported by Maynert and Van Dyke (155) on the per cent unchanged

$$\log \frac{1}{C} = -0.438\ (\log P)^2 + 1.579 \log P + 1.926$$
$$n = 13,\ r = 0.969,\ s = 0.098 \quad \ldots\ldots \quad 96$$

$$\log \%\ \text{excreted} = -1.235 \log P + 2.695$$
$$n = 10,\ r = 0.957,\ s = 0.224 \quad \ldots\ldots \quad 97$$

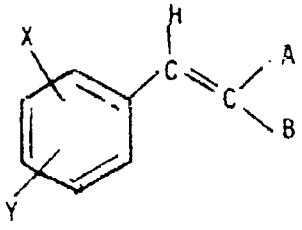
$$\log \%\ \text{metabolized } (\textit{in vitro}) = 0.511 \log P + 0.313$$
$$n = 4,\ r = 0.987,\ s = 0.063 \quad \ldots\ldots \quad 98$$

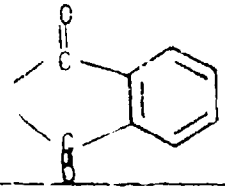
$$\log \%\ \text{metabolized } (\textit{in vivo}) = 0.634 \log P + 0.599$$
$$n = 3,\ r = 0.999,\ s = 0.026 \quad \ldots\ldots \quad 99$$

barbiturates excreted, Hansch *et al* (18) developed equation 97. Similarly from the data of Dorfman and Goldbaum (156) on the per cent of the barbiturates metabolized by liver (*in vitro*) and by mice (*in vivo*) Hansch *et al* (18) developed equations 98 and 99. In contrast to equation 96 equations 97 to 99 are linear with respect to log P but the slopes of these lines are in opposite directions. The optimum log P (that is log P_0) for equation 96 is 1.80 while the % excreted decreases with increasing lipophilicity (equation 97). The % of the barbiturates metabolized (equations 98 and 99), on the other hand, increases with increasing lipophilicity. The above conclusions are in accord with general observations. Hansch (44) states that "in general, the more water soluble a molecule is, the more rapidly it is eliminated in the urine". Brodie *et al* (157) have demonstrated the likelihood of a direct relationship between the rate of microsomal metabolism and the lipophilic character of drugs. Hansch *et al* (10,158) have since shown quantitatively that as members of a family of drugs become more lipophilic, other factors being equal, they are more rapidly destroyed by microsomal metabolism (48).

QSAR EQUATIONS INVOLVING EXPERIMENTALLY DETERMINED CONSTANTS ON MODEL SYSTEMS (84)

Holmes *et al* (84) obviated some of the above difficulties by using physico-chemical properties derived from *in vitro* model systems to evaluate the three dominant *in vivo* properties governing the degree of stimulatory activity of the conjugated heteroenoid compounds, I, on the frog flexor reflex. The rate of penetration of the I compounds to the receptor site was evaluated



A	B	A	B
COR	H	$COCH_3$	$COCH_3$
CO_2R	H	$COCH_3$	$CO_2C_2H_5$
NO_2	H	$CO_2C_2H_5$	$CO_2C_2H_5$
NO_2	R	$CO_2C_2H_5$	$CONH_2$
CN	CN	$CONH_2$	$CONH_2$
CO_2R	CN		
$CONH_2$	CN		
$CONR_2$	CN		

I

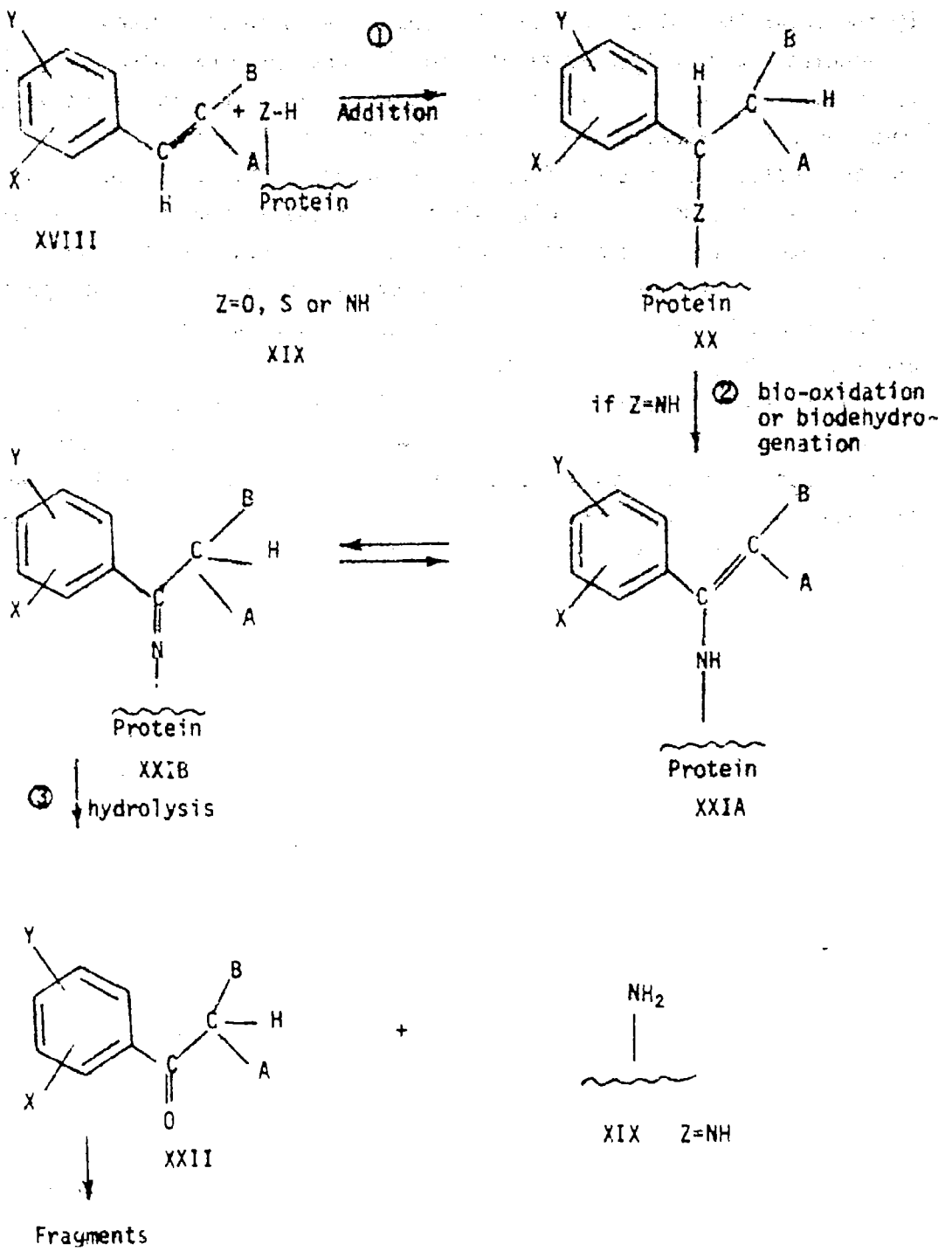
by the logarithms of their partition coefficients in the systems cyclohexane-water (P'), 1-octanol-water (P) and cyclohexanol-water (P"), while the rates of reaction of the I compounds with some component of the receptor were evaluated by *in vitro* second order rate constants, k, for the addition of nucleophiles to the conjugated system of the I compounds. The rate of wastage in the bulk medium and of the I compounds on the way to the receptor site was evaluated by the *in vitro* pseudo first order rate constants for the hydrolysis of these compounds by a reverse aldol process in Sörensen pH7 buffer, k_{H_2O}, and in bacterial growth medium II of Schmidt and Moyer (159), k_M. The stimulatory activity, A_T (in moles per litre) is the threshold concentration necessary for stimulation of the frog flexor reflex. The stimulatory activity, A_T, was considered to be proportional to the product of some power of the above three physico-chemical properties. This can be expressed mathematically by equation 100 or by its logarithmic form, equation 101. Any bio-catalytic effect that is present in the

$$A_T = t(P^a \times k^b \times k_{H_2O}^c) \quad \ldots \quad 100$$

$$\log A_T = a \log P + b \log k + c \log k_{H_2O} + \log t \quad \ldots \quad 101$$

in vivo reactions of the I compounds with the receptor site would be absent in the *in vitro* reactions of the model compound (mirroring the group or groups involved in reaction at the receptor site) with the unsaturated compound. This must be accommodated in equations 100 and 101.

A three step hypothesis (Fig. I) was advanced for the chemical reactions involved in the stimulation of the frog flexor reflex. Step number 1 involves the addition (XVIII → XX) of some receptor-nucleophile or -nucleophiles, XIX, to the conjugated system of the I compounds. This addition step just generates the substrate for bio-oxidation or bio-dehydrogenation (XX → XXIA ⇄ XXIB) in step number 2. If the receptor site is regenerated, then hydrolysis of XXIB to XXII (step number 3) regenerates the receptor-nucleophile, XIX, and XXII which, in turn, may suffer further degradation to smaller fragments or chelate with some metal. It was considered that the bio-oxidation or the bio-dehydrogenation (step 2)



Reactions Envisaged in the Stimulation of Frog Flexor Reflex.

FIG. I

induces stimulation or initiates the step that triggers the chain of reaction leading to stimulation.

If a reducing group were built right into the stimulatory molecule, then neither the double bond in the stimulatory molecule nor the addition step (step 1) in Fig. I would be required for stimulation. Since it is well known that catechol monomethyl ethers are readily oxidized to biphenols (160-164), then stimulation by these compounds may stem from the dimeric bio-oxidation or bio-dehydrogenation of these catechol monomethyl ethers to biphenols. Morphine and its isomers are readily converted, in the same way, to pseudo morphine (see ref. 84 page 9) and its isomers (165,166), so by analogy it was suggested that this might provide an insight into the chemical reactions associated with the biological properties of morphine and related compounds.

Introduction of the bio-oxidation or bio-dehydrogenation step would appear to raise the number of dominant factors from three to four for non-specific receptor sites. If the bio-oxidant at the receptor site, however, were a very strong oxidizing agent that would oxidize or dehydrogenate even weak reducing groups in the drug-substrate conjugate, then bio-oxidation would not be a stimulatory controlling step even if it is the step inducing stimulation. Equation 101 was applied to the stimulation of the non-specific receptor of the frog flexor reflex by the I compounds and then to the catechol monomethyl ethers. This equation was then successfully applied to the biological activities of the morphine alkaloids, the receptors for which are stereo- and structurally-specific.

The I compounds were demonstrated to inhibit the growth of four species of bacteria when incubated at 37°C for 17 hours in bacterial growth medium II of Schmidt and Moyer (159). Combining equations for the stimulatory activities of the I compounds on the frog flexor reflex with their growth inhibitory activities against bacteria permitted the calculation of stimulatory activities on the frog flexor reflex from the observed growth inhibitory activities against bacteria.

Equations for Stimulation of the Frog Flexor Reflex and for Inhibition of Growth of Microorganisms

Rates of hydrolysis of the I compounds with two angular A and B groups were so slow in Sörensen pH7 buffer and in 2% ethanol -98% water that it was considered that wastage in the time frame of the determination of A_T would be negligible so "c" of equation 101 was set at zero. Equations were developed by multiple linear regression analysis relating log A_T to log P, log P', log P", log k_{SH} (the second order rate constant for the *in vitro* addition of n-butanethiol to the I compounds) and k_{2SH} (the second order rate constant for the *in vitro* addition of 1,3-propanedithiol to the I compounds). These relationships are expressed in equations 102 to 105.

$$\log A_T = -0.21 \log P' - 0.41 \log k_{SH} - 4.59$$
$$n = 46,\ r = 0.96,\ s = 0.19 \quad \ldots\ldots \quad 102$$

$$\log A_T = -0.19 \log P' - 0.50 \log k_{2SH} - 4.13$$
$$n = 12,\ r = 0.95,\ s = 0.17 \quad \ldots\ldots \quad 103$$

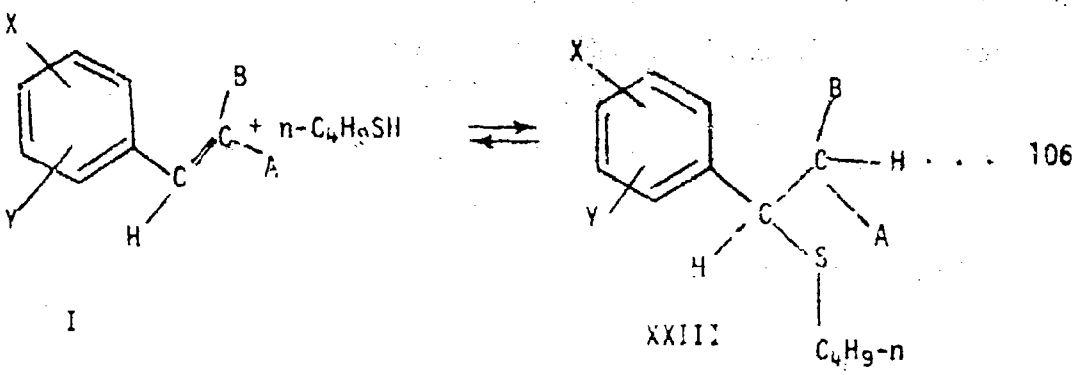
$$\log A_T = -0.42 \log P - 0.38 \log k_{SH} - 4.00$$
$$n = 27,\ r = 0.96,\ s = 0.19 \quad \ldots\ldots \quad 104$$

$$\log A_T = -0.42 \log P'' - 0.42 \log k_{SH} - 3.97$$
$$n = 9,\ r = 0.99,\ s = 0.10 \quad \ldots\ldots \quad 105$$

The similarity in the coefficient of log P' in equations 102 and 103 and of the analogous coefficients in equations 104 and 105 prompts the suggestion that these figures are characteristic for the biological medium through which the I compounds must pass. For convenience it might be termed the index of penetrability.

Second order rate constants, k_{SH}, for the addition of n-butanethiol to the I compounds have been determined as have the pseudo first order rate constants, k_{rev}, for the reverse reaction in equation 106. Equilibrium constants, K_{SH}, for this reaction were then derived from equation 107. A satisfactory equation 108 was developed by multiple linear regression relating log A_T to log P and log K_{SH} for the 3- and 4-substituted I compounds (A=B=$COCH_3$) but it would not accommodate the 2-substituted derivatives nor

would it accommodate diethyl benzalmalonate, I (A=B=$CO_2C_2H_5$).



. . . 106

$$K_{SH} = \frac{k_{SH}}{k_{rev}} \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad 107$$

$$\log A_T = -0.26 \log P' - 0.43 \log K_{SH} - 2.81$$
$$n = 6,\ r = 0.94 \quad \ldots\ldots\ldots\ldots \quad 108$$

The rate constants, k_{rev}, for the adducts of the I compounds (A=B=$COCH_3$) and of (A=B=$CO_2C_2H_5$) reveal that k_{rev} is markedly dependent upon steric factors of the A and B groups and of the substituent at C_2. It was also seen that, exclusive of the ortho- substituted-3-benzal-2,4-pentanediones, I(A=B=$COCH_3$), the rate constants, k_{rev}, are sensibly constant. For these compounds, then, the equilibrium constant, K_{SH}, will be a function of the forward rate constant, k_{SH}. This relationship is expressed in equation 109.

$$\log K_{SH} = 0.91 \log k_{SH} + 3.89$$
$$n = 6,\ r = 0.95 \quad \ldots\ldots\ldots\ldots\ldots \quad 109$$

An additional complication was considered when equations for the stimulation of the frog flexor reflex by the β-nitrostyrenes, I(A=NO_2, B=H, CH_3, C_2H_5), were being developed. The observation of Wallace *et al* (167) that under certain specified conditions nitro groups quantitatively oxidize thiols to disulphides, necessitated

the consideration of this as a potential source of wastage of the β-nitrostyrenes. This reductive wastage was evaluated by their polarographic half-wave potentials, $E_{\frac{1}{2}}$. Using the A component of the rate constant for the addition of $n\text{-}C_4H_9NH_2$ to the β-nitrostyrenes afforded equations 110-113.

$$\log A_T = -0.18 \log P' + 0.20 \log A - 4.88$$
$$n = 29,\ r = 0.74 \quad \ldots\ldots\ldots\ldots \quad 110$$

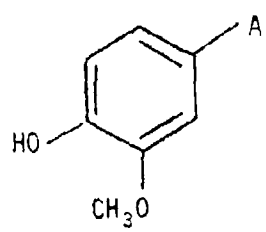
$$\log A_T = -0.18 \log P' + 0.21 \log A - 0.23 \log |E_{\frac{1}{2}}| - 4.94$$
$$n = 29,\ r = 0.74 \quad \ldots\ldots\ldots\ldots \quad 111$$

$$\log A_T = -0.72 \log P + 0.24 \log A - 3.62$$
$$n = 9,\ r = 0.83 \quad \ldots\ldots\ldots\ldots \quad 112$$

$$\log A_T = -0.76 \log P + 0.07 \log A + 3.74 \log |E_{\frac{1}{2}}| - 2.52$$
$$n = 9,\ r = 0.88 \quad \ldots\ldots\ldots\ldots \quad 113$$

Since there is no distinct improvement in the correlation coefficients when a $\log |E_{\frac{1}{2}}|$ term is introduced, it must be concluded that reductive wastage of the β-nitrostyrenes is not serious in the stimulation of the frog flexor reflex.

The three-step hypothesis for stimulation, Fig. I, suggests that a term involving rate constants is not necessary for a series of catechol monomethyl ethers, XXIV, with about the same critical oxidation potential (158,169,170). Equations 114 and 115 were developed by regression analysis.



A=R, CH_2CH_2COR, $CH_2CH_2CO_2R$

XXIV

$$\log A_T = -0.38 \log P' - 2.81$$
$$n = 15,\ r = 0.88 \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad 114$$

$$\log A_T = 0.09(\log P')^2 - 0.72 \log P' - 2.60$$
$$n = 15,\ r = 0.92 \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad 115$$

The experimental conditions for determining A_T values for these weakly active compounds were somewhat different from those used for the I compounds which may account for the different coefficient of log P'.

The chemical reactivity of the I compounds with angular A and B groups places these compounds in the central portion of the reactivity spectrum of the conjugated heteroenoid compounds. With the very sluggish members of this series the reactivity parameter becomes the stimulatory controlling step and the coefficient of log P or log P' approaches zero. At the other end of the spectrum, where the compounds are chemically very reactive, the partition parameter becomes the stimulatory controlling step and the coefficient of the rate parameter approaches zero. For the very reactive compounds, such as the 2-benzal-1,3-indanediones, the rate of hydrolytic wastage by a reverse aldol process becomes an important factor. As the rate of cyanide addition, k_{CN}, increases so does the rate of hydrolysis, k_{H_2O}, of the benzalmalononitriles, I(A=B=CN), in Sörensen pH7 buffer. This relationship is expressed in equation 116. As a result the -0.80 log k_{CN} term of equation 117 for the stimulatory activities of the benzalmalononitriles may be the algebraic sum of a stimulatory component and

$$\log k_{H_2O} = 1.09 \log k_{CN} - 3.03$$
$$n = 11,\ r = 0.96 \quad \ldots\ldots\ldots\ldots\ldots \quad 116$$

$$\log A_T = -0.26 \log P' - 0.60 \log k_{CN} - 4.77$$
$$n = 11,\ r = 0.86 \quad \ldots\ldots\ldots\ldots\ldots \quad 117$$

a wastage component with the stimulatory component being the larger of the two. The reverse appears to be true for 2-benzal-1,3-indanediones (ref. 84, p. 366).

The I compounds inhibit the growth of the bacteria *S. aureus*, *S.albus*, *E. coli* and *Aerobacter aerogenes* but in this review emphasis will be placed on their inhibitory activities against *S. aureus*. The incubation at 37°C was for 17 or 18 hours in bacterial growth medium II of Schmidt and Moyer (159),and not a matter of 5 minutes at 25°C in 2% ethanol -98% water as in the determination of A_T. Furthermore, Yasnikov and Gaivoronskaya (171) and Williamson and Witten (172) have demonstrated that amino acids and proteins catalyze the hydrolysis of diethyl benzmalonate, I(A=B=$CO_2C_2H_5$), and related compounds by a reverse aldol process so that it is not surprising to find that the rates of hydrolysis of these compounds in bacterial growth medium II of Schmidt and Moyer become significant in the equations for inhibition of growth, IG_{50}^{17} and IG_{50}^{18} when the times are 17 and 18 hours. The rates of hydrolysis were evaluated by the pseudo first order rate constants, k_W, for the hydrolysis of the I compounds in bacterial growth medium II. Regression analysis led to equations 118-120 for the I compounds with angular A and B groups.

For *Staphylococcus aureus*

$$\log IG_{50}^{17} = -0.24 \log P' - 0.55 \log k_{SH} + 1.07 \log k_W + 0.72$$
$$n = 52,\ r = 0.88 \quad \ldots\ldots\ldots \quad 118$$

$$\log IG_{50}^{17} = -0.51 \log P - 0.55 \log k_{SH} + 1.06 \log k_W + 1.44$$
$$n = 40,\ r = 0.86 \quad \ldots\ldots\ldots \quad 119$$

$$\log IG_{50}^{17} = -0.17 \log P' - 0.56 \log k_{2SH} + 0.73 \log k_W - 0.22$$
$$n = 11,\ r = 0.86 \quad \ldots\ldots\ldots \quad 120$$

Again the coefficients of log P' in equations 118 and 120 are very similar and furthermore the coefficients in equations 118 and 119 are the same except for those of the partition terms. The first two terms on the right hand side of equations 118-120 must be associated with the inhibitory process since,as log P, log P', log k_{SH} and log k_{2SH} become larger in a positive sense,log IG_{50}^{17} becomes larger (more active) in a negative sense. The third term on the right hand side of these equations must be associated with wastage since,as log k_W becomes larger in a positive sense,log IG_{50}^{17} becomes smaller (less active) in a negative sense.

Comparing equation 121 for the inhibition of growth IG_{50}^{17} (*S.alb*) of *S. albus* by the I compounds with two A and B activating groups with equation 118 for *S. aureus* reveals a change in the coefficients of all three parameters.

$$\log IG_{50}^{17}(S.alb) = -0,15 \log P' - 0.44 \log k_{SH} + 0.67 \log k_W - 0.75$$
$$n = 27,\ r = 0.70 \ldots\ldots \quad 121$$

The β-nitrostyrenes and their β-alkyl derivatives were examined to study the following factors.

1) Do dipole moments of the compounds evaluate the rates of penetration to the site of action as well as or better than partition coefficients?

2) Is reductive wastage of the nitro compound (see page 57) significant in the 17 hour period of incubation?

3) What effect has time of incubation upon the equation for inhibition of growth of *S. aureus*?

4) What effect has the addition of 1% albumin upon the equation for inhibition of growth of *S. aureus*?

Equations were developed from the data for *S. aureus* using for the rate parameter the A and B components for the rate constant for the addition of $n\text{-}C_4H_9NH_2$ to the I compounds (A=NO_2, B=H, CH_3, C_2H_5) and the rate constants, k_{SH}, for the addition of $n\text{-}C_4H_9SH$ to these compounds. A log $|E_s|$ term was then introduced into these equations and its effect upon the correlation coefficient examined. The log P' term was then replaced by the logarithm of the dipole moment μ and then combined with log P' to determine the effect upon the correlation coefficient. This is summarized in equations 122 to 133.

$$\log IG_{50}^{17} = -0.30 \log P' + 0.07 \log A + 0.70 \log k_W - 1.73$$
$$n = 35\ r = 0.90 \ldots\ldots\ldots \quad 122$$

$$\log IG_{50}^{17} = -0.28 \log P' - 0.08 \log A + 3.44 \log |E_{\frac{1}{2}}| + 0.64 \log k_W - 1.11$$
$$n = 35,\ r = 0.91 \quad \ldots\ldots \quad 123$$

$$\log IG_{50}^{17} = -0.26 \log P' + 0.16 \log B + 0.93 \log k_W - 0.77$$
$$n = 34,\ r = 0.91 \quad \ldots\ldots \quad 124$$

$$\log IG_{50}^{17} = -0.25 \log P' + 0.13 \log B + 1.10 \log |E_{\frac{1}{2}}| + 0.93 \log k_W - 0.52$$
$$n = 34,\ r = 0.91 \quad \ldots\ldots \quad 125$$

$$\log IG_{50}^{17} = -0.30 \log P' - 0.26 \log k_{SH} + 0.87 \log k_W - 0.60$$
$$n = 34,\ r = 0.91 \quad \ldots\ldots \quad 126$$

$$\log IG_{50}^{17} = -0.28 \log P' - 0.16 \log k_{SH} + 2.15 \log |E_{\frac{1}{2}}| + 0.90 \log k_W - 0.17$$
$$n = 34,\ r = 0.92 \quad \ldots\ldots \quad 127$$

$$\log IG_{50}^{17} = -0.27 \log P' + 3.00 \log |E_{\frac{1}{2}}| + 0.77 \log k_W - 0.74$$
$$n = 34,\ r = 0.91 \quad \ldots\ldots \quad 128$$

$$\log IG_{50}^{17} = 1.16 \log \mu - 0.10 \log A + 5.15 \log |E_{\frac{1}{2}}| + 1.01 \log k_W - 0.51$$
$$n = 12,\ r = 0.93 \quad \ldots\ldots \quad 129$$

$$\log IG_{50}^{17} = 1.13 \log \mu - 0.02 \log B + 5.30 \log |E_{\frac{1}{2}}| + 1.13 \log k_W - 0.01$$
$$n = 12,\ r = 0.93 \quad \ldots\ldots \quad 130$$

$$\log IG_{50}^{17} = 1.19 \log \mu + 0.235 \log k_{SH} + 5.62 \log |E_{\frac{1}{2}}| + 0.95 \log k_W - 1.09$$
$$n = 12,\ r = 0.93 \quad \ldots\ldots \quad 131$$

$$\log IG_{50}^{17} = 1.15 \log \mu + 5.07 \log |E_{\frac{1}{2}}| + 1.16 \log k_W - 0.04$$
$$n = 12,\ r = 0.93 \quad \ldots\ldots \quad 132$$

$$\log IG_{50}^{17} = -0.54 \log P' + 0.21 \log \mu - 0.51 \log k_{SH} + 2.94 \log |E_{\frac{1}{2}}| + 1.17 \log k_W + 2.13$$
$$n = 12,\ r = 0.97 \quad \ldots\ldots \quad 133$$

The procedure for the growth inhibitory activities, IG_{50}^{18}, reported by Schales and Graefe (173) for the β-nitrostyrenes and their β-alkyl derivatives is but a time modification of that used above. The time of incubation of the *S. aureus* was 18 hours at 37°C. Unfortunately data on all the compounds used above were not available for the derivation of equations 134 to 138.

$$\log IG_{50}^{18} = -0.23 \log P' + 0.19 \log A + 0.90 \log k_W - 0.71$$
$$n = 17,\ r = 0.94 \quad \ldots\ldots\ldots \quad 134$$

$$\log IG_{50}^{18} = -0.19 \log P' - 0.02 \log A + 9.62 \log |E_{\frac{1}{2}}| + 1.06 \log k_W + 2.29$$
$$n = 17,\ r = 0.97 \quad \ldots\ldots\ldots \quad 135$$

$$\log IG_{50}^{18} = -0.25 \log P' - 0.19 \log k_{SH} + 0.82 \log k_W - 0.56$$
$$n = 17,\ r = 0.94 \quad \ldots\ldots\ldots \quad 136$$

$$\log IG_{50}^{18} = -0.18 \log P' + 0.19 \log k_{SH} + 10.84 \log |E_{\frac{1}{2}}| + 0.92 \log k_W + 1.66$$
$$n = 17,\ r = 0.97 \quad \ldots\ldots\ldots \quad 137$$

$$\log IG_{50}^{18} = -0.18 \log P' + 9.43 \log |E_{\frac{1}{2}}| + 1.09 \log k_W + 2.34$$
$$n = 17,\ r = 0.97 \quad \ldots\ldots\ldots \quad 138$$

The average coefficient of log P' for equations 134 to 138 is -0.21 which does not differ much from the average coefficient of log P' -0.28 for equations 122-128. In general, the coefficient of log k_W and the proportionality constant are larger, in a positive sense, in equations 134-138 than they are in equations 122-128. These two factors contribute, in large part, to the decrease in growth inhibitory activities of these compounds when the time of incubation is increased from 17 to 18 hours. Schales and Graefe (173) found that the addition of 1% albumin to the bacterial growth medium II of Schmidt and Moyer reduced the growth inhibitory activities, IG_{50}^{18+a}, of the β-nitrostyrenes and their β-alkyl derivatives against *S. aureus*. *In vitro* studies* revealed that addition of 1% albumin had no significant effect upon the values for log P', log k_{SH} and log $|E_{\frac{1}{2}}|$. This amount of albumin, however, did catalyze the *in vitro* rates of hydrolysis of the β-nitrostyrenes slightly, while the catalytic effect upon the β-nitropropenylbenzenes was quite pronounced (see ref. 84 page 1165). Rate constants, k_{W+a}, for the hydrolysis of these compounds in medium II fortified with 1% albumin are reported in ref. 84 page 1165. Substituting log k_{W+a} for log k_W gave equations 139 to 141 for log IG_{50}^{18+a} against *S. aureus*.

---

* C.E. Lough. Unpublished data.

$$\log IG_{50}^{18+a} = -0.16 \log P' - 0.33 \log A + 11.51 \log |E_{\frac{1}{2}}| - 0.24 \log k_{W+a} - 2.10$$

$$n = 17,\ r = 0.81 \quad \ldots\ldots \quad 139$$

$$\log IG_{50}^{18+a} = -0.12 \log P' + 0.39 \log k_{SH} + 11.81 \log |E_{\frac{1}{2}}| + 0.23 \log k_{W+a} - 1.24$$

$$n = 17,\ r = 0.80 \quad \ldots\ldots \quad 140$$

$$\log IG_{50}^{18+a} = -0.18 \log P' + 3.82 \log |E_{\frac{1}{2}}| + 0.21 \log k_{W+a} - 2.43$$

$$n = 17,\ r = 0.70 \quad \ldots\ldots \quad 141$$

If the albumin is adsorbed on the surface of the bacteria or permeates into the interstices of the cell wall, then this will alter the permeability characteristics of the cell wall. These characteristics will trend towards those of the gram negative bacteria such as *E. coli*, with high protein content in the cell wall. The coefficients of log P' in equations 139-141 are more positive than those in equations 134-138 as are those in the equations for *E. coli*.

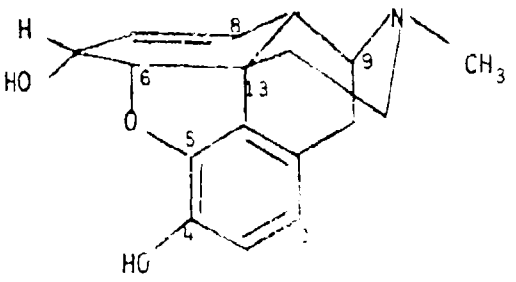
The antimicrobial action of the β-nitrostyrenes and their β-alkyl derivatives against *S. albus*, *E. coli*, *Aerobacter aerogenes*, *Pseudomonas aeruginosa*, *Trichophyton asteroidee*, *Botrytis allii*, *B. cinerea*, *Venturia inaequalis*, *Fusarium bulbigenum*, *Penicillium* Lumber Mould, *Aspergillus niger*, *Candida albicans* and *Trichomonas vaginalis* were treated similarly in ref. 84.

Equations for Various Biological Activities of the Morphine Alkaloids

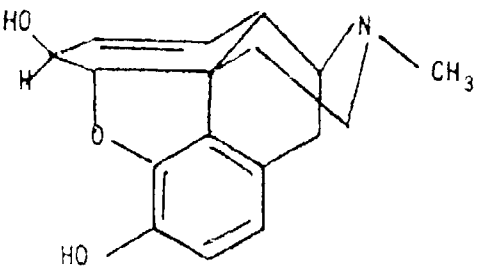
α-Isomorphine, XXVI, β-isomorphine, XXVII, and γ-isomorphine, XXVIII, are isomers of morphine, XXV. Morphine*, XXV, and α-isomorphine, are diastereomers as are β-and γ-isomorphines.

---

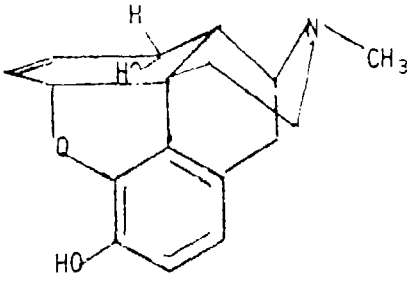
* For the structures of the morphine alkaloids see the treatises by Manske and Holmes (165) and by Bentley (177).



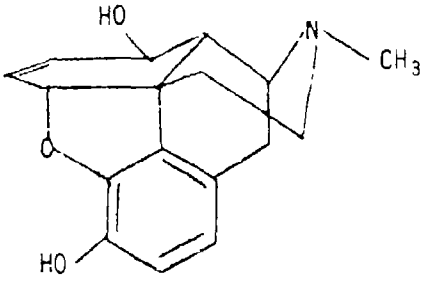
XXV



XXVI



XXVII



XXVIII

g-Isomorphine, XXVII, and morphine, XXV, are structural isomers, as are γ-isomorphine, XXVIII, and α-isomorphine, XXVI. Since the receptors involved in the biological actions are stereo- and structurally-specific (ref 84 page 460), a judicious choice of derivatives must be made from the large group (129 compounds) of morphine alkaloids for which biological activities have been reported by Kreuger, Eddy and Sumwalt (174). Analgesic, A_{analg}, excitant, A_{excit}, and emetic, A_{emet}, activities have been determined against cats, Ca; respiratory depressant, $A_{resp\ depr}$, activities have been determined against rabbits, Ra; while general depressant, $A_{gen\ depr}$, activities have been determined against rats, R. As well convulsant, A_{conv}, activities and lethal doses, $A_{L.D}$, have been determined on mice, M. More recently,

Drs L.J. Sargent and A.E. Jacobson* have determined the analgesic activities of 20 morphines and codeines on mice by the standard and reproducible (175) method of Eddy and Leimbach (176). The results from this method have the added advantage that all the data have been subjected to probit analysis.

On the hypothesis outlined on page 53, these biological activities should be dependent upon four dominant factors, 1) the rate of penetration to the receptor site, 2) the fit factor of the alkaloids with the receptor, 3) the rate of reaction at the receptor site and 4) the rate of wastage or of transformation to another morphine derivative prior to reaching the receptor site.

Since structural- and/or stereo-specific receptors are involved, the selection of morphine alkaloids and derivatives to be included in the regression analysis is much more difficult than for non-specific receptor sites and involves a number of arbitrary decisions as to which compounds have structures that can be considered to be sensibly the same. Any transformation of the morphine molecule, such as acetylation of the C_6-hydroxyl to 6-acetylmorphine or methylation of the C_3-hydroxyl group to codeine, represents a change of structure. Catalytic hydrogenation of the ethenoid linkage probably alters the conformation of morphine and its derivatives which, in turn, may alter the relationship between the rate of penetration and the logarithm of the partition coefficient. These changes will also influence the fit factors of the alkaloids relative to the topography of the receptor which, in turn, may alter the rates of reaction at the receptor site. As a result, many compromises must be accepted if a series of alkaloids is to be subjected to regression analyses. This and the larger error in the determination of the biological activities probably contribute to the poorer correlation coefficients for the plot of calculated activities against observed values.

---

* L.J. Sargent and A.E. Jacobson, National Institutes of Health, Bethesda, Maryland, 20014. Private Communication.

The five criteria used in the selection of the alkaloids for analysis were as follows:

1) the structure of the alkaloid must be firmly established;

2) the relationship between the rate of penetration to the receptor site and the logarithm of the partition coefficient, P''' (in ether-water), must be constant for all members of the series;

3) the fit factor of the alkaloids with respect to the topography of the receptor site must remain sensibly constant;

4) the mechanism of the chemical reactions involved at the receptor causing the biological response must be the same for all members of the series;

5) the rate of wastage of the members of the series must be sensibly constant.

These restrictions limited the study to an analysis of simple derivatives of morphine, α-isomers and their dihydro derivatives.

Equations first order and second order with respect to log P''' were developed by regression analysis for the biological tests reported by Kreuger, Eddy and Sumwalt and by Sargent and Jacobson. This analysis is summarized in equations 142 to 153a.

$$\log A_{analg\text{-}M}^{*} = -0.64 \log P''' - 5.67 \qquad n = 9,\ r = 0.95 \quad (142)$$

$$\log A_{analg\text{-}M} = -0.067 (\log P''')^2 - 0.62 \log P''' - 5.63 \qquad n = 9,\ r = 0.96 \quad (143)$$

$$\log A_{analg\text{-}Ca} = -0.35 \log P''' - 5.99 \qquad n = 14,\ r = 0.72 \quad (144)$$

---

* See list of symbols and abbreviations.

$$\log A_{analg\text{-}Ca} = 0.005 (\log P'''')^2 - 0.36 \log P'''' - 5.99$$
$$n = 14,\ r = 0.72 \quad \ldots\ldots \quad 145$$

$$\log A_{excit\text{-}Ca} = -0.57 \log P'''' - 5.96$$
$$n = 14,\ r = 0.86 \quad \ldots\ldots \quad 146$$

$$\log A_{excit\text{-}Ca} = 0.216 (\log P'''')^2 - 0.61 \log P'''' - 6.04$$
$$n = 14,\ r = 0.88 \quad \ldots\ldots \quad 147$$

$$\log A_{resp\ depr\text{-}Ra} = -0.81 \log P'''' - 6.95$$
$$n = 9,\ r = 0.92 \quad \ldots\ldots \quad 148$$

$$\log A_{resp\ depr\text{-}Ra} = -0.062 (\log P'''')^2 - 0.83 \log P'''' - 6.94$$
$$n = 9,\ r = 0.92 \quad \ldots\ldots \quad 149$$

$$\log A_{gen\ depr\text{-}R} = -1.04 \log P'''' - 5.10$$
$$n = 14,\ r = 0.94 \quad \ldots\ldots \quad 150$$

$$\log A_{gen\ depr\text{-}R} = 0.276 (\log P'''')^2 - 1.09 \log P'''' - 5.21$$
$$n = 14,\ r = 0.95 \quad \ldots\ldots \quad 151$$

$$\log A_{conv\text{-}M} = -0.88 \log P'''' - 3.19$$
$$n = 6,\ r = 0.96 \quad \ldots\ldots \quad 152$$

$$\log A_{LD\text{-}M} = -0.79 \log P'''' - 3.12$$
$$n = 6,\ r = 0.95 \quad \ldots\ldots \quad 153$$

$$\log A_{LD\text{-}M} = 0.129 (\log P'''')^2 - 0.84 \log P'''' - 3.18$$
$$n = 6,\ r = 0.96 \quad \ldots\ldots \quad 153a$$

This would indicate that the relationship between the logarithm of the partition coefficient and the rate of penetration holds for an iterated process involving transfer of a drug through many cells and across a multiplicity of membranes. Secondly, with judicious caution, this method for examining drug action can be applied to systems that involve stereo- and structurally-specific receptors.

Calculation of Biological Activities on One Organism from Observed Activities on Another

Equations have been developed relating the biological activities, A (in moles per litre), of the agonists, I, to several of their *in vitro* physico-chemical properties. These equations may be represented by the generalized equation 154 where O is some *in vitro* measure of the agonist as an oxidizing agent. Similar equations have been developed relating the bacterial growth inhibitory activities, IG (in moles per litre), of these agonists to the same *in vitro* physico-chemical properties. These equations can be represented by the generalized equation 155. In these equations the coefficients of some of the terms may be zero. Subtraction of equation 155 from 154 leads to equation 156. If the correlation coefficient for log IG (calc) vs log IG (obs) is good, then log IG (obs) can be substituted for log IG (calc) in equation 156 to give equation 157.

$$\log A\,(\text{calc}) = a \log P' + b \log k + c \log O + d \log k_{wast} + \log t \quad \ldots \quad 154$$

$$\log IG\,(\text{calc}) = a' \log P' + b' \log k + c' \log O + d' \log k_{wast} + \log t' \quad \ldots \quad 155$$

$$\log A\,(\text{calc}) = \log IG\,(\text{calc}) + (a-a') \log P' + (b-b') \log k + (c-c') \log O + (d-d') \log k_{wast} + \log t/t' \quad . \quad 156$$

$$\log A\,(\text{calc}) = \log IG\,(\text{obs}) + (a-a') \log P' + (b-b') \log k + (c-c') \log O + (d-d') \log k_{wast} + \log t/t' \quad . \quad 157$$

The correlation coefficient for log A (calc) vs log A (obs) for equation 157 would not be expected to be as high as that for equation 154 since two experimental errors are introduced, namely in the determination of log A (obs) and log IG (obs). In some cases within a single series of compounds the two experimental errors may deviate in the same direction from the true value, while in others, the deviations may be in opposite directions. A second factor that may lead to reduced correlation coefficients is if one term such as c log O in equation 154 is the algebraic sum of two terms representing two biological processes

such as stimulation and wastage, while the analogous term in equation 155 represents but a single process. Collinearity between terms in the equation will also add complications. Equation 157 has been applied to the data for a number of organisms using I compounds with angular A and B groups and I compounds where A = NO_2 and B = H, CH_3 and C_2H_5. This method applies equally well to the calculation of biological activities of the morphine alkaloids from their observed activities upon another. This is illustrated by equations 158 to 175 (Fig. II).

The analgesic activities of a few morphine and codeine derivatives on man have been reported (175). Equations have been developed from these limited data and the ultimate goal of calculating the biological activities of drugs on man from their observed activities on test animals appears now to be within reach.

Eqns. used in Derivation	Derived Equation	Eqn. No.
	I Compounds with Angular A and B Groups	
102 - 118	$\log A_T = \log IG_{50}^{17}$ (S.a.) + 0.03 log P' + 0.14 log k_{SH} -1.07 log k_W -5.32 n = 33, r = 0.93	158
103 - 120	$\log A_T = \log IG_{50}^{17}$ (S.a.) -0.02 log P' + 0.06 log k_{2SH} -0.73 log k_W -3.91 n = 9, r = 0.98	159
104 - 119	$\log A_T = \log IG_{50}^{17}$ (S.a.) + 0.10 log P + 0.17 log k_{SH} -1.06 log k_W -5.44 n = 27, r = 0.93	160
	I Compounds with A = NO_2, B = H, CH_3 and C_2H_5	
111 - 123	$\log A_T = \log IG_{50}^{17}$ (S.a.) + 0.10 log P' + 0.29 log A -3.67 log $\lvert E_{\frac{1}{2}} \rvert$ -0.64 log k_W -3.83 n = 21, r = 0.57	161
110 - 128	$\log A_T = \log IG_{50}^{17}$ (S.a.) + 0.09 log P' + 0.20 log A - 3.00 log $\lvert E_{\frac{1}{2}} \rvert$ -0.77 log k_W -4.14 n = 21, r = 0.59	162
	The Morphine Alkaloids	
146 - 144	$\log A_{excit.-Ca^*} = \log A_{analg.-Ca}$ - 0.22 log P''' + 0.03 n = 14, r = 0.88	163
147 - 145	$\log A_{excit.-Ca} = \log A_{analg.-Ca}$ + 0.211 (log P''')^2 -0.26 log P''' -0.06 n = 14, r = 0.90	164
148 - 144	$\log A_{resp.depr.-Ra} = \log A_{analg.-Ca}$ -0.46 log P''' -0.98 n = 9, r = 0.92	165

Eqns used in Derivation	Derived Equation	Eqn No.
149 - 145	$\log A_{\text{resp. depr.-Ra}} = \log A_{\text{analg.-Ca}} - 0.067 (\log P''')^2 - 0.48 \log P''' - 1.00$	
	$n = 9,\ r = 0.94$	166
150 - 144	$\log A_{\text{gen.depr.-R}} = \log A_{\text{analg.Ca}} - 0.69 \log P''' + 0.89$	
	$n = 14,\ r = 0.94$.	167
151 - 145	$\log A_{\text{gen.depr.-R}} = \log A_{\text{analg.-Ca}} + 0.271 (\log P''')^2 - 0.74 \log P''' + 0.78$	
	$n = 14,\ r = 0.96$.	168
152 - 144	$\log A_{\text{conv.-M}} = \log A_{\text{analg.-Ca}} - 0.48 \log P''' + 2.80$	
	$n = 6,\ r = 0.99$	169
153 - 144	$\log A_{\text{LD-M}} = \log A_{\text{analg.-Ca}} - 0.35 \log P''' + 2.96$	
	$n = 6,\ r = 0.83$	170
153a- 145	$\log A_{\text{LD-M}} = \log A_{\text{analg.-Ca}} + 0.124 (\log P''')^2 - 0.49 \log P''' + 2.81$	
	$n = 6,\ r = 0.83$	171
148 - 146	$\log A_{\text{resp.depr.-Ra}} = \log A_{\text{excit.-Ca}} - 0.24 \log P''' - 1.00$	
	$n = 9,\ r = 0.86$	172
149 - 147	$\log A_{\text{resp.depr.-Ra}} = \log A_{\text{excit.-Ca}} - 0.278 (\log P''')^2 - 0.22 \log P''' - 0.94$	
	$n = 9.\ r = 0.86$	173
150 - 146	$\log A_{\text{gen.depr.-R}} = \log A_{\text{excit.-Ca}} - 0.47 \log P''' + 0.86$	
	$n = 14,\ r = 0.95$	174

Eqns used in Derivation	Derived Equation	Eqn. No.
151 - 147	$\log A_{gen.depr.-R} = \log A_{excit.-Ca} + 0.060 (\log P''')^2 - 0.48 \log P''' + 0.84$ $n = 14, r = 0.95$	175

* See list of symbols and abbreviations.

Equations for Calculating Biological Activities of Some Morphine Alkaloids on One Organism from the Observed Activities on Another Organism.

FIG. II

Summary

The locus of points for the activities of a series of drugs against one organism, when plotted against those of another organism, are often far from linear and the reason for this becomes apparent from a comparison of log A_T values for the I compounds with angular A and B groups with those for the stimulation, A_{BM}**(in moles per litre)of the mouse eye. The relation of observed log A_{BM} and log A_T values is given in equation 176. To determine the extent, if any, of wastage in the

$$\log A_{BM} = 0.62 \log A_T - 0.72$$
$$n = 13,\ r = 0.51 \quad \ldots\ldots\ldots\ldots \quad 176$$

stimulatory process on the frog flexor reflex by the I compounds with angular A and B groups, the pseudo first order rate constant, k_W, for hydrolysis in bacterial growth medium and log $|E_{\frac{1}{2}}|$ were inserted in equation 102 to give equation 177. Equation 178 is the analogous equation for stimulation, A_{BM}, of the mouse eye.

$$\log A_T = -0.20 \log P' - 0.45 \log k_{SH} - 1.47 \log |E_{\frac{1}{2}}| - 0.03 \log k_W - 4.61$$
$$n = 22,\ r = 0.98 \quad \ldots\ldots\ldots\ldots \quad 177$$

$$\log A_{BM} = -0.05 \log P' + 0.05 \log k_{SH} + 1.96 \log |E_{\frac{1}{2}}| - 0.78 \log k_W - 6.74$$
$$n = 13,\ r = 0.77 \quad \ldots\ldots\ldots\ldots \quad 178$$

These equations are quite different. In both cases the coefficient of the log k_W term is negative so it is associated with stimulation* rather than wastage. Furthermore the coefficient of log P' in equation 177 is much larger in a negative sense than that in equation 178. Subtraction of equation 177 from 178 and replacement of log A_T (calc) by log A_T (obs) (r=0.98) yields equation 179. The correlation coefficient for the plot of

$$\log A_{BM} = \log A_T \text{ (obs)} + 0.15 \log P' + 0.50 \log k_{SH} + 3.43 \log |E_{\frac{1}{2}}| - 0.75 \log k_W - 2.14$$
$$n = 13,\ r = 0.92 \quad \ldots\ldots\ldots\ldots \quad 179$$

---

* Log k_W in equation 178 is a better model than log k_{SH} for evaluating the rate of reaction at the receptor site.

** These values determined by sequential blepharospasm test by Mr. B.J. Wenner.

log A_{BM} (calc) vs log A_{BM} (obs) is high (0.98). Comparing equation 179 with 176 it will be seen that the extra terms in equation 179 account for the improved correlation coefficient. The larger coefficient of log P' in a negative sense in equation 177 compared to that in equation 178 indicates that the frog's leg favours the penetration of lipophilic (oil-soluble) compounds to a greater extent than does the mouse eye. The equations for the β-nitrostyrenes and their β-alkyl derivatives against *S. aureus* and *B. allii* clearly illustrate this point. Equations 180 to 183 were developed from the data of McGowan *et al* (178) for log IG_{100}. The average coefficient

$$\log IG_{100} = 0.58 \log P' - 0.02 \log A - 0.82 \log k_W - 8.67$$
$$n = 6,\ r = 1.00 \quad \ldots\ldots\ldots\ldots \quad 180$$

$$\log IG_{100} = 0.60 \log P' - 0.02 \log A + 5.32 \log |E_1| - 0.35 \log k_W - 5.56$$
$$n = 19,\ r = 1.00 \quad \ldots\ldots\ldots\ldots \quad 181$$

$$\log IG_{100} = 0.59 \log P' - 0.15 \log k_{SH} + 6.78 \log |E_2| + 0.03 \log k_W - 3.48$$
$$n = 22,\ r = 1.00 \quad \ldots\ldots\ldots\ldots \quad 182$$

$$\log IG_{100} = 0.60 \log P' + 5.32 \log |E_2| - 5.46$$
$$n = 19,\ r = 1.00 \quad \ldots\ldots\ldots\ldots \quad 183$$

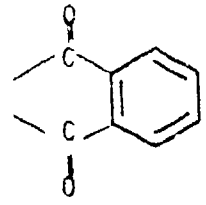
of log P' for equations 122 to 128 related to *S. aureus* is -0.28 while that for equation 180-183 for *B. allii* is + 0.59. The log IG_{100} value for 3-methoxy-4-hydroxy-β-nitrostyrene (log IG_{100} = -5.39) against *B. allii* reveals that it is 10 times more active than β-nitrostyrene (log IG_{100} = -4.38) even though it is chemically much less reactive than β-nitrostyrene. This difference in activity stems mainly from the relatively greater hydrophilic property of 3-methoxy-4-hydroxy-β-nitrostyrene (log P' = + 0.04) compared to that of β-nitrostyrene (log P' = + 1.80). The high lipophilic property (log P' = +3.45) of β-nitrobutenylbenzene, I(A=NO_2, B=C_2H_5), contributes materially to the low activity (log IG_{100} = -2.85) of this compound.

For those equations involving a hydrolytic wastage term to hold for the I compounds, the activities of the two derived fragments

must be either very low or a constant value. The log IG_{50}^{17} values of the 3- and 4-substituted-benzaldehydes against *S. aureus* were one order of magnitude lower than for the I compounds with two angular A and B groups and their log IG_{50}^{17} values were sensibly constant. The activities of the other hydrolytic fragments were low.

Methods have been developed for calculating log P, log P' (page 19), log k_{SH} and log k_W purely from structure on the blackboard which permits the calculation of log A_T and log IG_{50}^{17} values purely from their chemical structures.

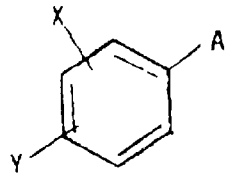
Covariance or collinearity between the various *in vitro* parameters is a complication that makes interpretation of the mechanism of the *in vivo* process difficult. One such covariance is expressed mathematically in equation 116. So any substituent introduced into the phenyl group of the I compounds to enhance reactivity will also enhance the rate of wastage. Maximum stimulatory activity of the I compounds with A=B=CN or

AB=  will result, then, from a happy compromise between

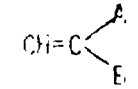
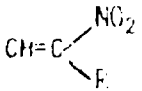
enhanced activity due to increased rate of reaction on the one hand and loss of activity due to an increased rate of wastage on the other. This optimum can be approached from either of the two extremes (84 page 393).

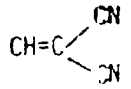
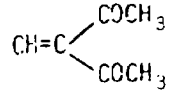
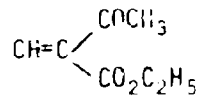
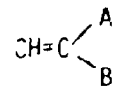
1) Introduce stronger activating groups at A and B in the I compounds to increase the rate of reaction.
2) Introduce large steric groups into the 2 and 6 position of the 2-benzal-1,3-indanediones to slow down the rate of reaction.

The parameters used in these equations have been plotted against each other to test for collinearity or covariance. These equations are catalogued in Fig. III.



A	Equation	n	Corr. Coeff.	Eqn. No.
CHO	$\log \mu = -0.26 \log P' + 0.81$	4	0.96	184
$CH=C(NO_2)R$	$\log \mu = 0.33 \log P' - 0.35$	6	0.71	185
$CH=C(A)CN$	$\log \mu = -0.08 \log P' + 0.81$	10	0.74	186
$CH=C(COCH_3)COCH_3$	$\log k_{SH} = 0.51 \log P' - 0.31$	6	0.56	187
$CH=C(COCH_3)CO_2C_2H_5$	$\log k_{SH} = 0.44 \log P' - 0.74$	10	0.65	188
$CH=C(CO_2C_2H_5)CO_2C_2H_5$	$\log k_{SH} = 0.37 \log P' - 2.12$	13	0.72	189
$CH=C(CONH_2)CONH_2$	$\log k_{SH} = 0.17 \log P' - 1.32$	8	0.17	190

A	Equation	n	Corr. Coeff.	Eqn. No.
$CH{=}CH_2NO_2$	$\log k_W = -0.28 \log A - 3.77$	18	0.67	191
 all angular A & B groups	$\log k_W = 0.46 \log k_{SH} - 3.84$	26	0.90	192
CHO	$\log \lvert E_{\frac{1}{2}} \rvert = -0.037 \log P' - 0.054$	10	0.79	193
	$\log \lvert E_{\frac{1}{2}} \rvert = -0.037 \log k_{CN} - 0.387$	11	0.83	194
	$\log \lvert E_{\frac{1}{2}} \rvert = -0.180 \log \mu - 0.179$	4	0.92	195
	$\log \lvert E_{\frac{1}{2}} \rvert = 0.006 \log P' - 0.232$	37	0.15	196
	$\log \lvert E_{\frac{1}{2}} \rvert = 0.033 \log A - 0.263$	35	0.77	197
	$\log \lvert E_{\frac{1}{2}} \rvert = -0.049 \log k_{SH} - 0.163$	26	0.78	198
	$\log \lvert E_{\frac{1}{2}} \rvert = 0.177 \log \mu - 0.319$	19	0.81	199
	$\log \lvert E_{\frac{1}{2}} \rvert = -0.086 \log k_W - 0.602$	26	0.75	200

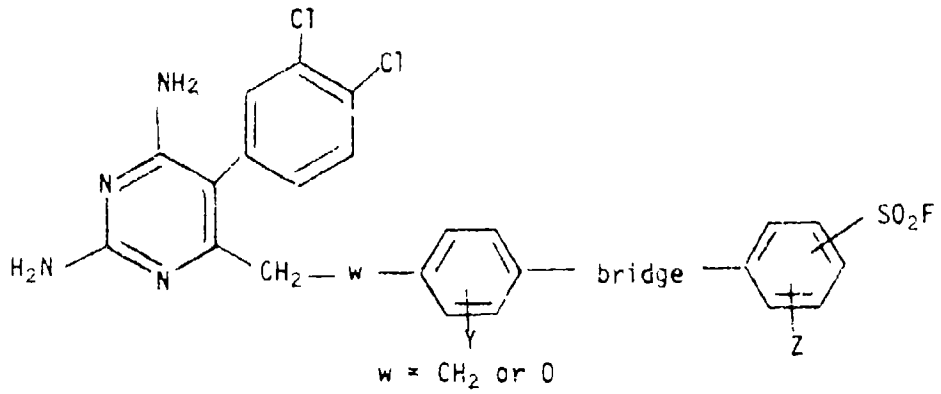
A	Equation	n	Corr. Coeff.	Eqn. No.
	$\log \lvert E_{\frac{1}{2}} \rvert = -0.026 \log P' + 0.095$	13	0.30	201
	$\log \lvert E_{\frac{1}{2}} \rvert = -0.080 \log AK + 0.106$	13	0.77	202
	$\log \lvert E_{\frac{1}{2}} \rvert = -0.152 \log k_{CN} + 0.002$	11	0.92	203
	$\log \lvert E_{\frac{1}{2}} \rvert = -0.131 \log k_{H_2O} - 0.391$	11	0.89	204
	$\log \lvert E_{\frac{1}{2}} \rvert = -0.037 \log P' + 0.066$	8	0.78	205
	$\log \lvert E_{\frac{1}{2}} \rvert = 0.077 \log P - 1.197$	8	0.70	206
	$\log \lvert E_{\frac{1}{2}} \rvert = 0.017 \log k_{SH} - 1.148$	13	0.32	207
 all angular A & B groups	$\log \lvert E_{\frac{1}{2}} \rvert = -0.060 \log k_{SH} + 0.046$	22	0.89	208
	$\log \lvert E_{\frac{1}{2}} \rvert = -0.043 \log k_W - 0.120$	17	0.12	209

Equations Reflecting Collinearity Between Various Physico-Chemical Properties.

FIG. III

DE NOVO SUBSTITUENT CONSTANTS

A completely empirical quantitative approach to structure-activity relationships was first presented in 1956 by Bruice *et al* (93) and later taken up by Free and Wilson (94), Ban and Fujita (179), Beasley and Purcell (180), Craig (181) and Clayton and Purcell (182). The basic assumption in this method is that substituent effects (in biological units) when added to the biological activity of the parent compound give the biological activity of the derivative. If the biological activities of a series of drugs belonging to one family are available a series of simultaneous equations can be set up with one unknown for each substituent in a given position. Solution of n simultaneous equations for n unknowns provides substituent constants in biological units for the series of drugs in question and for that specific biological process. Craig used 69 2-phenyl quinolines to yield 69 equations for their antimalarial activity. The substituent constants, so derived, are listed in references 181 and 44. Hansch, Silipo and Stellier (65) applied this method to the study of the inhibition of dihydrofolate reductase by 105 2,4-diamino-5(3,4-dichlorophenyl)-6-substituted pyrimidines, XXIX. A correlation coefficient of 0.920 was obtained for these 105 compounds.



XXIX

The advantage of this approach is that all properties of each substituent (hydrophobic, electronic, steric and other more subtle properties) are encompassed in each substituent constant. Modern chemistry has long recognized the great importance of electronic and steric factors of substituents on rate and equilibrium processes. As a result it would seem highly unlikely that constants derived for one set of drugs on one biological process would be applicable to another set of drugs applied to a different biological process. Group interactions complicate the situation still further.

DISCUSSION OF RESULTS

Hansch clearly demonstrated that the degree of biological activity of drugs is dependent upon three dominant factors:

1) hydrophobic effects,
2) electronic effects,
3) steric effects.

This method, however, takes into account neither wastage by metabolism and/or elimination nor the effects of hydrogen bonding. The method of Holmes takes the rate of wastage and hydrogen bonding into account but it has not been extended yet to include steric factors. The Free-Wilson method lumps all factors into one substituent constant.

The rates of penetration or the hydrophobic effects are most important in determining the degree of biological activity of drugs. For equation 102, log P' accounts for 73% of the variance in the variable log A_T. Hansch has shown that the partition coefficient is associated with, and evaluates two effects.

1) The non-specific effect which evaluates the random walk through the biological medium to the site of action (41).
2) The specific effect which is a measure of the weak hydrophobic bonding between drug and enzyme (48).

The non-specific role of the hydrophobic effect is dramatically illustrated by the equation derived by Hansch *et al* (11) of the quantitative determination by Solway (181) of the amount, C_b, of mono- and di-substituted benzeneboronic acids reaching the brain of mice 15 minutes after injection.

By regression analysis of the data, Hansch *et al* (11) showed that for penetration into the brain the best fit is given by equation 210.

$$\log C_b = -0.540\,\pi^2 + 0.765\,\pi + 1.505$$
$$n = 25,\ r = 0.915,\ s = 0.214 \quad \ldots\ldots\ldots \quad 210$$

Inclusion of σ in the regression analysis had no effect. In this case π evaluates the random walk of the benzeneboronic acids to the brain.

Many examples of the specific hydrophobic effect are known. Many equations have been developed by Hansch relating log P or π to the binding of drugs to pure enzymes. These equations are generally first order with respect to log P or π as is the case for binding molecules of 42 organic compounds to bovine serum albumin (44). Equation 211 relates the concentration, C, necessary for the formation of a 1:1 complex to log P

$$\log \frac{1}{C} = 0.75 \log P + 2.30$$
$$n = 42,\ r = 0.960,\ s = 0.159 \quad \ldots\ldots\ldots \quad 211$$

The series of 42 compounds is not a homologous series but is comprised of phenols, anilines, aliphatic alcohols, ketones and naphthalene as well as rigid molecules such as hydroxyadamantane and camphorquinone. Molecules which hydrogen bond well, such as phenols and alcohols, are accommodated equally as well by equation 211 as are molecules with little or no ability to hydrogen bond such as naphthalene, azobenzene and chloronitrobenzene and those prone to form charge transfer complexes all fit the above equation. Hansch and Anderson (19) presented evidence that hydrophobic bonding between a long hydrocarbon side chain and an aromatic ring contributes to maintaining such a molecule in a folded form thus materially altering its log P value. Hydrophobic bonding is considered to be an integral part of the partitioning process (20). Blanketing of part of a molecule by large bulky groups or by virtue of part of a molecule being held in a rigid position can prevent solvation of parts of the molecule thus leading to anomalies. The following two examples clearly illustrate this point. Blanketing has also been observed by Currie *et al* (113,84 pages 96 and 106).

$$\pi_{i\text{-}C_3H_7} - \pi_{n\text{-}C_3H_7} = -0.13$$

$$\pi_{t\text{-}C_4H_9} - \pi_{n\text{-}C_4H_9} = -0.22$$

Log P_0 and π_0

It has been demonstrated many times that equations for the biological activities of drugs in living organisms (the benzeneboronic acids in mice; equation 210) are second order with respect to log P, while *in vitro* studies with enzymes often lead to equations that are linear with respect to this parameter. In many instances this has led to failure in the design of inhibitors for living animals from the data for optimum activity for *in vitro* inhibition of enzymes. Baker found many extremely potent inhibitors (*in vitro*) for dihydrofolate reductase which were quite ineffective on leukemia cells. The locus of the points in the plot of a second order equation with respect to log P (equation 210) is a parabola, while that for a first order equation with respect to log P (211) is a straight line. Equation 211 mathematically states that binding power increases indefinitely with increase in log P, while in the case of equation 210 an optimum concentration is reached at a certain π value and then the concentration reaching the brain progressively falls off with increase in log P. The optimum value of π, namely π_0, for equation 210 is obtained by taking the derivative* of log C_b with respect to π and setting this equal to zero. Log P for the parent benzeneboronic acid is about +1.6

$$\frac{d \log C_b}{d\pi} = -2 \times 0.54\,\pi + 0.765$$

$$2 \times -0.54\,\pi + 0.765 = 0$$

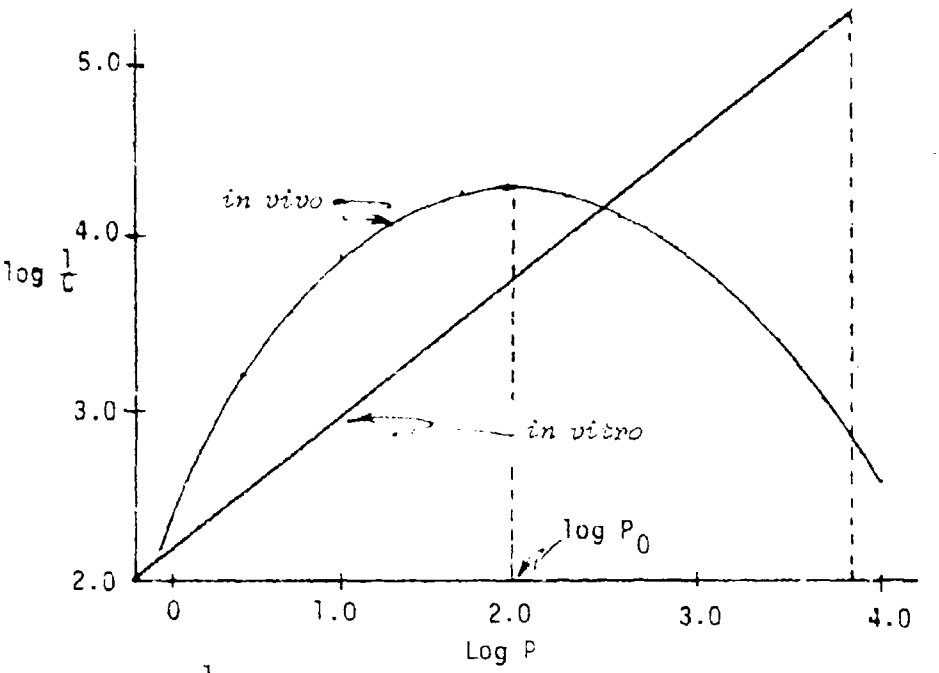
$$\pi_0 = 0.71$$

so log P_0 will be 1.6 + 0.71 = 2.31. The explanation of Baker's observations is shown graphically in Fig. IV. The activity of the *in vitro* inhibitor with log P = +3.90 would be very high (straight line) but the same compound would have a very low activity in the

---

* If c is included in the equation then the partial derivative with respect to π must be taken.

in vivo experiment (parabola)*.



Plot of log $\frac{1}{C}$ vs log P for an *in vitro* Inhibitor of an Enzyme and for the same Inhibitor in a Living Organism.

FIG. IV

* The *in vitro* action of antihistamines on guinea pig ileum (127) and the *in vivo* action of these compounds upon guinea pigs (128) is the exception to the above statement. Kutter and Hansch (27) have developed equations relating these two activities to substituent constants. Both activities are independent of π and π^2 terms and are dependent solely upon E_S. Hence the rate of penetration to the site of action in the *in vivo* experiments is not a controlling factor in determining the biological response. In fact both sets of data correlate very well on one equation.

Equations 212 to 227 derived by Hansch *et al* (18) from others' (182-189) data for the hypnotic activity of barbiturates and non-barbiturates permitted the calculation of log P_0 values (Fig. V). These data were obtained in a number of laboratories from 1923 to 1949. Not only was hypnosis defined in different ways such as ED and MED_{50} but some workers used rabbits, others mice and still others rats. Even the techniques used in the same laboratory varied over the years. The average log P_0 for equations 212 to 227 is + 1.98 and this does not differ markedly from that for the localization of the benzeneboronic acids (log P_0 = 2.31) in the brain of mice (equation 210). Similar results have been obtained for the inhibition of growth of Gram negative (equations 228-234) and Gram positive (equations 235 to 243) bacteria. The constancy of log P_0 values for each of these biological systems led Hansch *et al* (18) to conclude that "other factors being constant, sets of congeners (families of drugs) acting by the same mechanism on the same receptor sites should have the same log P_0 values." So if a new series of hypnotics is being developed then the member of the series with log P = ~ +2.0 should have the maximum hypnotic activity. While log P_0 or π_0 is a very useful function, it must be emphasized that it is an empirically determined constant and that other factors may influence the value of log P_0. For example side effects such as metabolism and elimination, which are log P-dependent, may materially influence the value of log P_0. For many non-polar functions π and molar volume (MR) are collinear so increase in π may influence the fit factor between drug and receptor.

It is surprising how many of the equations in Fig. V have about the same intercept (the average value of the intercept is + 1.686) in view of the different test animals used and the diversity of techniques employed. In certain instances intercepts can be very helpful in the design of drugs.

Test	Derived Equations	Log P_0	Eqn No.
AD_{50}(mice)	$\log \frac{1}{C} = -0.438 (\log P)^2 + 1.579 \log P + 1.926$ $n = 13, r = 0.969, s = 0.098$	1.80	212
MED(rabbits)	$\log \frac{1}{C} = -0.630 (\log P)^2 + 2.092 \log P + 1.918$ $n = 11, r = 0.896, s = 0.140$	1.66	213
MED(rabbits)	$\log \frac{1}{C} = -0.529 (\log P)^2 + 2.377 \log P + 1.351$ $n = 9, r = 0.744, s = 0.139$	2.25	214
MAD(rats)	$\log \frac{1}{C} = -0.173 (\log P)^2 + 0.719 \log P + 2.653$ $n = 17, r = 0.531, s = 0.099$	2.08	215
MED(rats)	$\log \frac{1}{C} = -0.545 (\log P)^2 + 1.804 \log P + 2.098$ $n = 15, r = 0.855, s = 0.124$	1.65	216
AD_{50}(mice)	$\log \frac{1}{C} = -0.690 (\log P)^2 + 2.797 \log P + 0.672$ $n = 13, r = 0.702, s = 0.219$	2.03	217
ND(mice)	$\log \frac{1}{C} = -0.236 (\log P)^2 + 1.273 \log P + 1.867$ $n = 10, r = 0.915, s = 0.132$	2.69	218
AD_{50}(mice)	$\log \frac{1}{C} = -0.240 (\log P)^2 + 1.300 \log P + 1.948$ $n = 14, r = 0.737, s = 0.914$	2.71	219
HD_{50}(mice)	$\log \frac{1}{C} = -0.219 (\log P)^2 + 0.864 \log P + 2.501$ $n = 6, r = 0.858, s = 0.178$	1.97	220

Test	Derived Equations	Log P_0	Eqn No.
HD_{50}(mice)	$\log \frac{1}{C} = -0.686\ (\log P)^2 + 2.451 \log P + 0.724$ $n = 8,\ r = 0.965,\ s = 0.058 \ldots$	1.79	221
HD_{50}(mice)	$\log \frac{1}{C} = -0.510\ (\log P)^2 + 2.134 \log P + 0.857$ $n = 8,\ r = 0.944,\ s = 0.105 \ldots$	2.09	222
HD_{50}(mice)	$\log \frac{1}{C} = -0.675\ (\log P)^2 + 2.099 \log P + 1.663$ $n = 8,\ r = 0.947,\ s = 0.082 \ldots$	1.56	223
MHD(rabbits)	$\log \frac{1}{C} = -0.231\ (\log P)^2 + 1.020 \log P + 1.515$ $n = 11,\ r = 0.826,\ s = 0.114 \ldots$	2.21	224
ED_{50}(guinea pigs)	$\log \frac{1}{C} = -0.414\ (\log P)^2 + 1.589 \log P + 1.322$ $n = 13,\ r = 0.805,\ s = 0.130 \ldots$	1.92	225
HD_{50}(mice)	$\log \frac{1}{C} = -0.314\ (\log P)^2 + 0.999 \log P + 1.983$ $n = 6,\ r = 0.913,\ s = 0.108 \ldots$	1.59	226
MED(mice)	$\log \frac{1}{C} = -0.177\ (\log P)^2 + 0.599 \log P + 1.893$ $n = 14,\ r = 0.918,\ s = 0.079 \ldots$	1.69	227

Equations and Log P_0 for Hypnosis by Barbiturates (Eqs 212-219) and non-Barbiturates (Eqs 220-227).

FIG. V

Bacteria	Derived Equations	Log P_0	Ref	Eq. No.
S. typhosa	$\log PC' = -0.280 (\log P)^2 + 2.199 \log P + 1.219 \sigma - 2.215$ $n = 11, r = 0.972, s = 0.169$	3.93	190	228
	$\log PC' = -0.180 (\log P)^2 + 1.628 \log P - 1.777$ $n = 11, r = 0.975, s = 0.208$	4.52	190	229
	$\log PC' = -0.204 (\log P)^2 + 1.771 \log P - 1.871$ $n = 10, r = 0.982, s = 0.180$	4.35	190	230
	$\log PC' = -0.407 (\log P)^2 + 3.082 \log P + 2.460 \sigma - 3.649$ $n = 12, r = 0.971, s = 0.168$	3.79	190	231
	$\log PC' = -0.334 (\log P)^2 + 2.991 \log P - 4.540$ $n = 26, r = 0.936, s = 0.190$	4.48	190	232
E. coli	$\log \frac{1}{C} = -1.040 (\log P)^2 + 8.531 \log P + 0.774 \sigma - 12.629$ $n = 9, r = 0.967, s = 0.138$	4.10	190	233
	$\log \frac{1}{C} = -0.226 (\log P)^2 + 2.088 \log P - 1.126$ $n = 19, r = 0.479, s = 0.438$	4.62	190	234

Equations and Log P_0 for Inhibition of Growth of Gram Negative Bacteria.

FIG. VI

Bacteria	Derived Equations	Log P_0	Ref.	Eq. No.
S. aureus	$\log \frac{1}{C} = -0.335 (\log P)^2 + 3.453 \log P + 2.995 \sigma - 4.200$ $n = 12, r = 0.899, s = 0.770$	5.15	190	235
	$\log PC' = -0.167 (\log P)^2 + 2.121 \log P - 3.498$ $n = 35, r = 0.961, s = 0.236$	6.36	190	236
	$\log PC' = -0.147 (\log P)^2 + 1.733 \log P - 2.211$ $n = 12, r = 0.995, s = 0.093$	5.90	190	237
	$\log PC' = -0.167 (\log P)^2 + 1.784 \log P - 2.201$ $n = 8, r = 0.996, s = 0.066$	5.34	190	238
	$\log \frac{1}{C} = -0.264 (\log P)^2 + 3.081 \log P - 4.416$ $n = 5, r = 0.991, s = 0.131$	5.84	190	239
Strep. viridans	$\log \frac{1}{C} = -0.247 (\log P)^2 + 2.815 \log P - 2.301$ $n = 5, r = 0.994, s = 0.094$	5.69	190	240
Strep. faecalis	$\log \frac{1}{C} = -0.125 (\log P)^2 + 1.359 \log P + 0.415$ $n = 10, r = 0.861, s = 0.334$	5.42	190	241
B. diphtheria	$\log \frac{1}{C} = -0.123 (\log P)^2 + 1.431 \log P + 1.161$ $n = 17, r = 0.936, s = 0.300$	5.81	190	242

Bacteria	Derived Equations	Log P_0	Ref.	Eq. No.
Cl. sporogenes	$\log \frac{1}{C} = -0.189 (\log P)^2 + 2.373 \log P - 2.631$ $n = 5, r = 0.985, s = 0.164$	6.27	190	243

Equations and Log P_0 for Inhibition of Growth of Gram Positive Bacteria.

FIG. VII

Slopes and Intercepts in the Equations

Comparison of slopes and intercepts for equations involving different families of compounds on the same biological test system or different families of compounds on different test systems provides an insight into the similarity of the biological processes and their sensitivity to response by these chemicals. The slopes of equations 244 to 264 are a measure of the sensitivity of the biological systems to the hydrophobicity of the chemical series examined. Equations 244-264 have been developed by Hansch *et al* (42) from data gleaned from the literature on the antimicrobial activities of various esters of p-hydroxybenzoic acid against a number of Gram positive and Gram negative organisms as well as some fungi. These equations are listed in Fig. VIII. The average slope of the Gram positive organisms is 0.863, while those for the Gram negative organisms and the fungi are respectively 0.540 and 0.518. There is a marked difference in the sensitivity of the Gram positive organisms to the hydrophobic properties of the esters of p-hydroxybenzoic acid from those of the other two groups. Within the Gram positive organism there is little constancy in the intercept.

The magnitude of the intercept is determined by the sensitivity of the biological system and the chemical reactivity of the active groups in the families of drugs under examination. If the same family of drugs is applied to two different biological systems then the intercept reflects the sensitivity of response of these two systems to this family of chemicals. Selection of an ester of p-hydroxybenzoic acid with a $P = 1$ or $\log P = 0$ and substitution of this value in equations 244 and 248 yields equations 265 and 266.

Organism	Type* of Org.	Derived Equations	n	r	s	Eq. No.
S. aureus	+	$\log \frac{1}{C} = 0.841 \log P - 1.011$	7	0.989	0.086	244
C. albicans	+	$\log \frac{1}{C} = 0.726 \log P + 0.496$	7	0.957	0.249	245
C. albicans	+	$\log \frac{1}{C} = 0.742 \log P + 0.469$	7	0.973	0.197	246
S. aureus	+	$\log \frac{1}{C} = 0.955 \log P - 0.248$	4	0.991	0.109	247
B. subtilis	+	$\log \frac{1}{C} = 0.848 \log P + 0.197$	4	0.993	0.084	248
B. cereus	+	$\log \frac{1}{C} = 1.072 \log P - 0.272$	4	0.982	0.175	249
S. lutea	+	$\log \frac{1}{C} = 0.955 \log P - 0.248$	4	0.991	0.109	250
S. cerevisiae	+	$\log \frac{1}{C} = 0.732 \log P + 0.661$	4	0.998	0.044	251
S. cerevisiae	+	$\log \frac{1}{C} = 0.848 \log P + 0.497$	4	0.993	0.084	252
S. pastorianus	+	$\log \frac{1}{C} = 0.848 \log P + 0.497$	4	0.993	0.084	253
P. vulgaris	-	$\log \frac{1}{C} = 0.457 \log P + 1.051$	4	0.956	0.120	254
K. pneumoniae	-	$\log \frac{1}{C} = 0.624 \log P + 0.966$	4	0.999	0.011	255
A. niger	F	$\log \frac{1}{C} = 0.417 \log P + 1.981$	12	0.975	0.081	256
P. roqueforti	F	$\log \frac{1}{C} = 0.508 \log P + 1.972$	7	0.994	0.064	257

Organism	Type of Org.	Derived Equation	n	r	s	Eq. No.
P. roqueforti	F	$\log \frac{1}{C} = 0.440 \log P + 1.862$	7	0.981	0.099	258
B. fulva	F	$\log \frac{1}{C} = 0.501 \log P + 1.819$	7	0.976	0.125	259
A. niger	F	$\log \frac{1}{C} = 0.502 \log P + 1.266$	5	0.921	0.148	260
R. nigricans	F	$\log \frac{1}{C} = 0.624 \log P + 1.266$	4	1.000	0.011	261
T. lignorum	F	$\log \frac{1}{C} = 0.396 \log P + 2.044$	4	0.964	0.093	262
T. mentagrophytes	F	$\log \frac{1}{C} = 0.622 \log P + 1.769$	4	0.999	0.014	263
T. rubrum	F	$\log \frac{1}{C} = 0.622 \log P + 1.769$	4	0.999	0.014	264

* The symbol + indicates gram positive organisms, - indicates gram negative organisms and F indicates fungi.

Linear Equations for Antimicrobial Activities of some Esters of p-Hydroxybenzoic Acid against Gram Positive, Gram Negative Organisms and Fungi (42).

FIG. VIII

S. aureus $\log \frac{1}{C} = -1.011$ 265

B. subtilis $\log \frac{1}{C} = +0.197$ 266

These two organisms have the same sensitivity to hydrophobicity, but the sensitivities (equations 265 and 266) of these two organisms to this chemical are quite different. This must be associated with the stereo-electronic character of the group in this ester that is involved in the inhibition of growth of these two organisms.

Collinearity of Parameters

Any significance attached to relationships discerned in multi-parameter equations of the form of type equations 4-10 (pages 36 to 50) must be tempered by a consideration of possible covariance between parameters*. Covariance between various *in vitro* physico-chemical properties used in determining log A_T and log IG_{50} values has been discussed on pages 76 to 79.

---

* Another factor which can lead to false conclusions is if the activities of metabolites are of the same order of magnitude as that of the drug. Usually metabolites are quite inactive, however, an exception to this is the primary metabolites of Δ^8- and Δ^9-tetrahydrocannabinol (192).

For certain collections of substituents Hansch has demonstrated that a correlation exists, to a degree, between π and σ (37,41), π and MR (58,60,61,67,70), σ^* and E_S (37) and MR and E_S (61). A correlation matrix was then presented by Hansch *et al* (99) for eight parameters involving 90 substituents.

APPLICATIONS OF QSAR

QSAR has not yet been developed and refined to the stage where it can forecast a "new find" or drug for some specific biological process but does provide a systematic and direct route to the development of the most effective drug in the drug family of the "new find". Examination of the equations developed for an already known series of drugs upon this biological process reveals the relative importance of hydrophobic electronic and steric factors and permits the calculation of the optimum log P_0 value and in many cases E_{S-0} values for the best drug in the "new find". The above factors also shed some light upon the topography of enzyme surfaces. Combining this information with a good knowledge of chemistry should lead to a series of specifications for a drug to fit the requirements for the specific biological process. It would seem to the author, then, but a short step to successful drug design.

Predictions for More Active Congeners

Log P_0 and signs of coefficients of the various terms in the generalized equations 267 and 268 provide an insight into the trends that will lead to optimum activity.

$$\log \frac{1}{C} = -\log C = a\pi + b\sigma + cE_S + \log t \quad \text{267}$$

$$\log IG_{50}^{17} = a' \log P + b' \log k_{SH} + c' \log k_W + \log t \quad \text{268}$$

If the coefficients "a" and "b" of equation 267 are positive then, as π and σ^{ϕ} get larger in a positive sense, log C will get larger in a negative sense (more active). If the coefficient "c" of equation 267 is positive then as E_S* gets larger in a negative sense log C

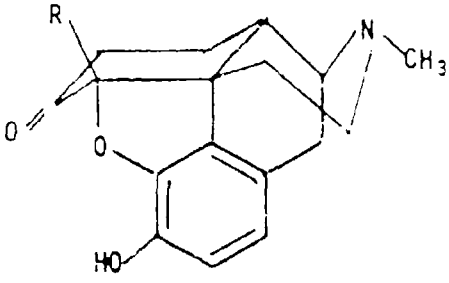
ϕ A positive sign associated with σ indicates that it is an electron -withdrawing substituent (CN, NO_2 etc) while a negative σ value implies that it is an electron donating substituent (CH_3, CH_3O, $(CH_3)_2N$ etc)

* As the substituent gets larger E_S becomes larger in a negative sense.

becomes larger in a positive sense (less active). So with positive coefficients in equation 267, increases in lipophilicity and stronger electron withdrawing groups will enhance biological activity, while a decrease in size of the substituent will produce a trend in the same direction. The signs of the coefficients in equation 268 reveal whether this term is associated with the biological process or with wastage (see page 60). Moreover the larger "a'" is in a negative sense the more the biological media favour the penetration of lipophilic compounds.

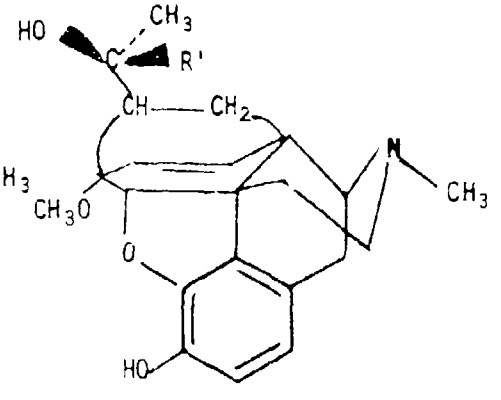
If the equation for the biological activity is second order with respect to log P or π, then log P_0 or π_0 can be calculated and a maximum value can be assigned to log P or π for maximum biological activity. Illustrative examples of this will be found in references 18, 21, 23, 35, 54, 58, 61, 66, 67, 68.

Analgesic and excitant activities on cats, general respiratory activities on rats and lethal doses on mice for a number of morphine alkaloids have already been calculated from equations 144-147, 150, 151, 153 and 153a. Equations 145, 147, 151 and 153a were used unsuccessfully to calculate the respective activities for six 5-alkyldihydromorphinones, XXX (ref 84 page 1425). As the R group of XXX gets larger, the deviation between



R=H, CH_3, C_2H_5, i-C_3H_7, n-C_5H_{11}, C_6H_5

XXX



XXXI

calculated and observed log A values increases, which suggests that steric interaction between the 5-alkyldihydromorphinone and the receptor plays a very important role in determining the biological activities of the members of this series. Attempts are now made in equations 269 to 276 to evaluate the steric effects by E_S from the data in Table 4. Correlation coefficients for log $A_{excit-Ca}$ and log A_{LD-M} are good while those for log $A_{analg-Ca}$ and log $A_{gen.depr.-R}$ are not.

$$\log A_{analg-Ca} = + 0.25 \log P^{m} + 0.41 E_S - 6.05$$
$$n = 6,\ r = 0.45 \quad (269)$$

$$\log A_{analg-Ca} = -1.24 (\log P^{m})^2 + 1.90 \log P^{m} + 1.18 E_S - 6.08$$
$$n = 6,\ r = 0.56 \quad (270)$$

$$\log A_{excit-Ca} = -0.25 \log P^{m} + 0.43 E_S - 6.27$$
$$n = 6,\ r = 0.82 \quad (271)$$

$$\log A_{excit-Ca} = -0.88 (\log P^{m})^2 + 0.92 \log P^{m} + 0.97 E_S - 6.29$$
$$n = 6,\ r = 0.85 \quad (272)$$

$$\log A_{gen.depr.-R} = -0.18 \log P^{m} + 0.36 E_S - 5.16$$
$$n = 6,\ r = 0.61 \quad (273)$$

$$\log A_{gen.depr.-R} = -1.40 (\log P^{m})^2 + 1.59 \log P^{m} + 1.89 E_S - 5.20$$
$$n = 6,\ r = 0.70 \quad (274)$$

$$\log A_{LD-M} = -0.40 \log P^{m} + 0.20 E_S - 4.01$$
$$n = 6,\ r = 0.98 \quad (275)$$

$$\log A_{LD-M} = -0.17 (\log P^{m})^2 - 0.18 \log P^{m} + 0.31 E_S - 4.01$$
$$n = 6,\ r = 0.98 \quad (276)$$

The sign of log P^{m} terms in equations 271, 273 and 275 is negative so greater lipophilicity leads to greater activity.

TABLE 4

Log P^{m} , E_S Log $A_{analg-Ca}$, Log $A_{excit-Ca}$, Log $A_{gen.depr-R}$, Log A_{LD-M} for some Alkyldihydromorphinones

Alkyl Group	Log P^{m}	E_S*	Log $A_{analg-Ca}$	Log $A_{excit-Ca}$	Log $A_{gen.depr-R}$	Log A_{LD-M}
H	-0.43	+1.24	-5.39	-5.54	-4.85	-3.54
CH_3	-0.10	0.00	-6.63	-6.48	-5.00	4.08
C_2H_5	+0.23	-0.07	-6.27	-6.54	-5.19	-4.06
i-C_3H_7	+0.48	-0.47	-5.39	-6.10	-5.16	-4.28
n-C_5H_{11}	+1.22	-0.40	-5.99	-6.35	-6.21	-4.51
C_6H_5	+1.37	(+0.23 (-2.58	-5.71	-6.44	-4.86	-4.58

* The values for E_S were taken from reference 44.

The $(\log P'')^2$ terms in equations 270, 272 and 274 reveal that increase in log P'' in a positive sense contributes to a larger log A term in a negative sense to a limit, and further increase leads to a decrease in activity (parabolic function). The positive coefficient of E_S indicates that as E_S becomes larger in a negative sense so does the log A term (more active). Increase in size of the R group of XXX favours increased activity. Introduction of an $(E_S)^2$ term into equations 269, 271 and 273 leads to equations 269a, 271a and 273a.

$$\log A_{analg\text{-}Ca} = +0.013 \log P'' + 1.75 (E_S)^2 - 1.22 E_S - 5.51$$
$$n = 6,\ r = 0.94 \quad \text{269a}$$

$$\log A_{excit\text{-}Ca} = -0.37 \log P'' + 0.89 (E_S)^2 - 0.40 E_S - 6.50$$
$$n = 6,\ r = 0.93 \quad \text{271a}$$

$$\log A_{gen.depr\text{-}R} = -0.11 \log P'' - 0.47 (E_S)^2 + 0.79 E_S - 5.04$$
$$n = 6,\ r = 0.65 \quad \text{273a}$$

Steric factors do not contribute materially to the degree of activity in $\log A_{LD\text{-}M}$ but they do in determining the magnitude of $\log A_{analg\text{-}Ca}$ and $\log A_{excit\text{-}Ca}$. $\log A_{analg\text{-}Ca}$ and $\log A_{excit\text{-}Ca}$ are parabolic functions of both E_S and log P''. As might be expected, then, for equation 269b where $(\log P'')^2$ and $(E_S)^2$ terms are inserted into equation 269, the correlation should be better than that for equation 269a. The correlation coefficient of 0.45 for equation 269 has now

$$\log A_{analg\text{-}Ca} = -1.13 (\log P'')^2 + 1.51 \log P + 1.71 (E_S)^2 - 0.48 E_S - 6.53$$
$$n = 6,\ r = 0.99 \quad \text{269b}$$

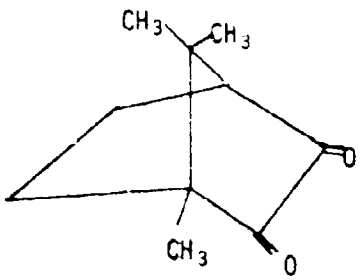
risen to 0.99. This emphasizes the importance of steric factors in determining the extent of analgesia by these compounds. Increase in the size of the group in the upper part of XXX favours increased activity on two counts (hydrophobic and steric) so it is not surprising to find that the analgesic activity of XXXI (R'=n-C_3H_7) is 12,000 times greater than that of morphine (193). There are two values for E_S for the phenyl group. The value of +0.23 refers to the thickness of the phenyl group while -2.58 refers to the breadth of the benzene ring.

The value of -2.58 gives completely irrational results so this gives some insight into the space requirements at the receptor sites.

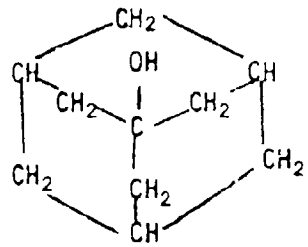
From equation 272, log P_0'' for this series of 5-alkyldihydromorphinones is +0.52. All other things being equal, then, maximum activity should be attained by the i-propyl homolog, but because of the large positive coefficient of E_S in equation 271 this factor exerts a pronounced influence in determining the magnitude of log $A_{excit-Ca}$. Log P_0'' for the lethal dose of these alkaloids on mice is -0.53.

Mapping of Enzyme Surfaces

The early model of drug-enzyme or drug-membrane interaction, namely the "lock and key" concept, served a useful role; however, recent developments have shown that protein structure and conformation is governed to a large extent by weak bonds such as hydrophobic bonding aided by hydrogen bonding and/or dipolar interactions. As a result the rigid "lock and key" concept has given way to one where enzymes have sites involving greater fluidity. This is supported by the fact that, at least over limited ranges, there is a linear relationship between interaction and log P (equations 40, 41 and 211) and E_S (equation 46). Such a relationship would hardly be expected on the "lock and key" hypothesis. The 42 compounds embodied in the data for equation 211 include simple alcohols, phenols and anilines as well as bulky molecules such as camphorquinone, XXXII, neopentyl alcohol and hydroxyadamantane, XXXIII.



XXXII

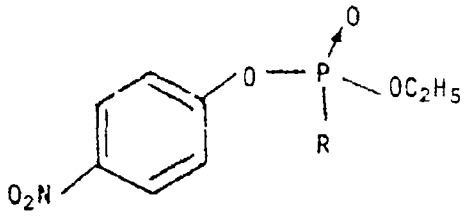


XXXIII

The fact that rigid bulky molecules fit equation 211 as well as flexible compounds argues against binding by lock and key interaction. While some conclusions on enzyme space were drawn prior to 1970, the majority of the work has appeared since 1973.

period 1966 - 1968

Equations 277 and 278 were developed by Hansch *et al* (37,15) from the data of Metcalf and Fukuto (125) for the substituent effects of R on the inhibition of cholinesterase by alkylphosphonic acid esters, XXXIV. Comparison of equation 278 with 277 reveals that neither



XXXIV

$$\text{Log } K = 3.74\ E_S + 7.54$$

$$n = 13,\ r = 0.901 \quad \ldots\ldots\ldots\ldots \quad 277$$

$$\log K = 0.15\ \pi - 1.68\ \sigma^* + 4.05\ E_S + 7.21$$

$$n = 13,\ r = 0.907 \quad \ldots\ldots\ldots\ldots \quad 278$$

hydrophobic (π) nor electronic (σ^*) factors operate in the inhibitory process. Thus it would appear that the phosphonate esters interact with the enzyme in such a way that contact is not made with hydrophobic regions of the enzyme.

Early in their classical work on the inhibition of dihydrofolate reductase by pyrimidines, Baker and Shapiro (194) concluded that inhibitors are bound to the enzyme by interaction of ring electrons with an electron-deficient site and by hydrophobic

interaction of one or more side chains. Equations 279 to 281 support these conclusions in a quantitative way (17).

$$\log \frac{1}{C} = -5.162\ \sigma - 5.002$$
$$n = 16,\ r = 0.760 \quad \ldots\ldots\ldots\ldots \quad 279$$

$$\log \frac{1}{C} = 0.302\ \pi - 1.970$$
$$n = 16,\ r = 0.328 \quad \ldots\ldots\ldots\ldots \quad 280$$

$$\log \frac{1}{C} = 0.457\ \pi - 5.820\ \sigma - 6.951$$
$$n = 16,\ r = 0.903 \quad \ldots\ldots\ldots\ldots \quad 281$$

Introduction of π^2, σ^2 or π^2 and σ^2 terms into these equations made no significant improvement in the correlation. The positive coefficient of π in equations 280 and 281 reveals that the more lipophilic the side chain the more effective the inhibitor. In like vein, the negative coefficient of σ requires that the more the substituent releases electrons to the ring the more effective the derivative is as an inhibitor. This is in accord with the conclusions of Baker and Shapiro.

By density gradient centrifugation, Fouts (195,196) separated the smooth-surfaced (s) particles from the denser rough-surfaced (r) particles of the endoplasmic reticulum of liver cells and determined the rate of metabolism of various drugs on both particles.* Fouts concluded that the difference in rates of metabolism by the two particles was due to different concentrations of the enzymes in the two particles.

Lien and Hansch (25) developed equations 282 and 283 for the ratio $R_{(r/s)}$ of enzyme activity of the two types of particles from rabbit liver microsomes prepared by the Rothschild and the Dallner methods.

---

* There was a difference in the rate of metabolism of drugs on the two particles from one to ten depending on the drug.

$$\log R_{(r/s)} = -0.101 \log P + 0.859$$
$$n = 10,\ r = 0.885,\ s = 0.116 \quad \ldots\ldots \quad 282$$

$$\log R_{(r/s)} = -0.091 \log P + 0.553$$
$$n = 7,\ r = 0.923,\ s = 0.089 \quad \ldots\ldots \quad 283$$

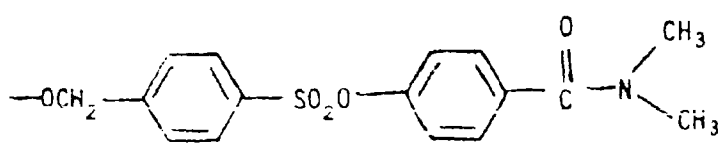
The above equations permit an alternative interpretation. The same concentration or activity of the enzymes can be present in both particles but the availability of the drug to the enzyme is rate limiting, and depends upon the environment in which the enzyme is set. From equation 282, the more lipophilic the drug the smaller the ratio in the rate of metabolism by the two particles. If some of the enzymes in each particle are in a very lipophilic surrounding this would explain why lipophilic drugs like benzpyrene are metabolized at a ratio near one. On this basis the smooth particles must have as well a set of enzymes in a hydrophilic medium which would account for the high ratio for drugs with a low log P value.

period 1974 to 1976

In 1974 Hansch and Silipo (58) turned their attention to Baker's expanding data on diamino-1,3,5-triazines, XVII, as inhibitors of dihydrofolate reductase. Eighty-three derivatives of XVII were included in the analysis with substituents, X*, at C_3, C_4 and C_3C_4. Early in the work it became apparent that there was a large difference in influence of hydrophobic properties of substituents at C_3 and C_4 as was the case for bulk evaluation by MR_3 and MR_4. Electronic effects evaluated by $\Sigma\sigma$ contributed nothing to improving the correlation. Linear combination of 10 parameters [π_3, π_4, $\pi_{3,4}$, MR_3, MR_4, $MR_{3,4}$, $\sigma_{3,4}$, $(\pi_3)^2$, $(\pi_4)^2$ and $(\pi_{3,4})^2$]produced 783 equations. The equation with two variables having the lowest standard deviation was 284 while the analogous three parameter equation was 285.

---

* These substituents ranged in size from H to



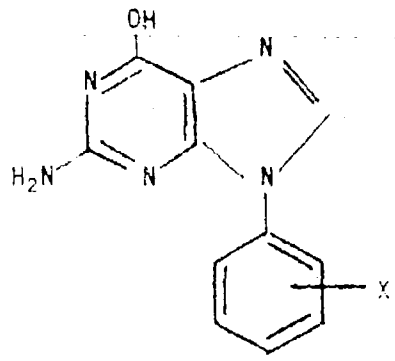
$$\log \frac{1}{C} = 0.464\ \pi_3 + 0.181\ MR_4 + 6.613$$
$$n = 83,\ r = 0.834,\ s = 0.422 \quad \ldots \quad 284$$

$$\log \frac{1}{C} = -0.127\ (\pi_3)^2 + 0.890\ \pi_3 + 0.150\ MR_4 + 6.618$$
$$n = 83,\ r = 0.905,\ s = 0.328 \quad \ldots \quad 285$$

Equation 285 accounts for 82% of the variance in $\log \frac{1}{C}$. Failure to account for 18% is not surprising considering the enormous variation in the size of substituents, X, in XVII.

Since σ plays no significant role in the equations, the electronic nature of the substituents, X, need not be considered. Furthermore π_4 and MR_3 do not contribute significantly to any of the equations. The reason why π_4 does not contribute materially to $\log \frac{1}{C}$ in these equations is that π (1-octanol-water) is not a good model for the hydrophobic bonding of the C_4 substituent (see page 16). From equations 284 and 285 Hansch (58) suggests that "there are two kinds of substituent space (meta and para) in or on the enzyme. Functions in the 3 position appear to be placed in a typical hydrophobic milieu. The coefficient of about 1 for this term is observed quite commonly. Substituents in the 4 position appear to be thrust into a more apolar region which π_4 does not model well. It seems likely that groups in the 4 position cause inhibition by producing conformational changes in dihydrofolate reductase by more firmly attaching the inhibitor through dispersion forces or by a combination of both." Adding an $(MR_4)^2$ term to the equations does not improve the correlation so $\log \frac{1}{C}$ is not a parabolic function of MR, so still bulkier substituents could be accommodated where C_4-substituents interact with the enzyme. On the other hand $(\pi_3)^2$ terms contribute materially to equation 285 so $\log \frac{1}{C}$ is a parabolic function of π and it is unwise to introduce a group at C_3 with a π value greater than π_0 which is +3.5. To examine the effect of bulkier groups at C_4 in the diamino-1,3,5-triazines, Silipo and Hansch (69) extended this study to include the analysis of the interaction of 244 of these compounds with dihydrofolate reductase.

Silipo and Hansch (66) have analyzed the data of Baker and Wood (197,198) on the inhibition, $\frac{1}{C}$, of xanthine oxidase by thirty 9-(substituted-phenyl) guanines, XXXV.



XXXV

These workers developed 511 equations involving 8 variables chosen from Π_2, Π_3, Π_4, $\Pi_{3,4}$, $(\Pi_{3,4})^2$, $\Sigma\Pi$, $(\Sigma\Pi)^2$ MR_2, MR_3, MR_4, $MR_{3,4}$, $(MR_{3,4})^2$ ΣMR, $(\Sigma MR)^2$, E_{S-2}, E_{S-4}, σ, σ^+, $\mathcal{F}_3$, $\mathcal{F}_4$, $\Sigma\mathcal{F}$, $\mathcal{R}_2$, $\mathcal{R}_3$, $\mathcal{R}_4$, $\Sigma\mathcal{R}$, D and D_2. In all, more than 2,000 equations were examined. The equation with the highest correlation coefficient and lowest standard deviation is equation 286. The relative importance of the various terms in equation 286 is manifest in the statistical data for equations 287 to 290.

$$\log \frac{1}{C} = 0.203\ MR_{3,4} + 1.259\ E_{S-2} + 0.432\ E_{S-4} + 4.327$$
$$n = 30,\ r = 0.924,\ s = 0.228 \quad . \quad 286$$

$$\log \frac{1}{C} = 0.24\ E_{S-4} + 6.12$$
$$n = 30,\ r = 0.359,\ s = 0.535 \quad . \quad 287$$

$$\log \frac{1}{C} = 0.26\ MR_{3,4} + 5.90$$
$$n = 30,\ r = 0.476,\ s = 0.505 \quad . \quad 288$$

$$\log \frac{1}{C} = 1.13\ E_{S-2} + 4.98$$
$$n = 30,\ r = 0.630,\ s = 0.445 \quad . \quad 289$$

$$\log \frac{1}{C} = 1.47\ E_{S-2} + 0.40\ E_{S-4}$$
$$n = 30,\ r = 0.857,\ s = 0.301 \quad . \quad 290$$

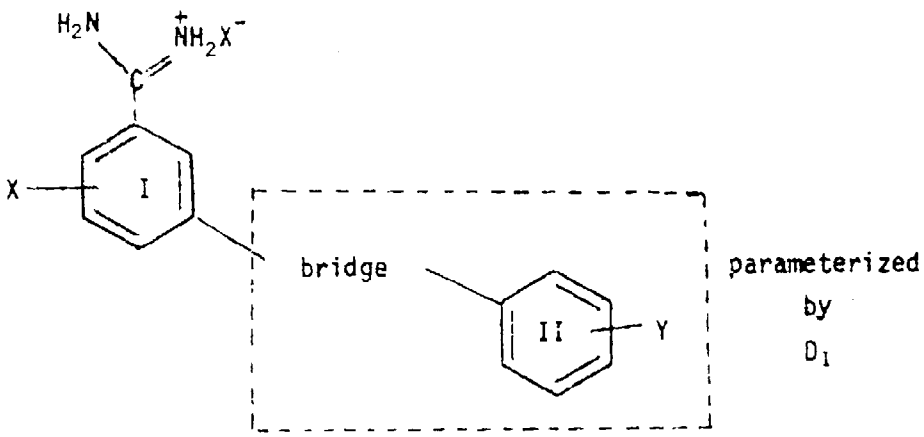
The outstanding observation to be made from equation 286 is that the 9-phenyl ring of these compounds does not locate itself in hydrophobic space on the enzyme. In no case did π give as good results as MR (see collinearity page 95). Equation 286 accommodates large substituents (e.g., 3-$NHCOC_6H_5$ and 4-$OCH_2CH_2CH_2C_6H_5$) in the meta and para position indicating the great flexibility in the enzyme space around these positions. Since inclusion of an $(MR_{3,4})^2$ term in the equation did not improve the correlation, then log $\frac{1}{C}$ does not appear to be a parabolic function of $MR_{3,4}$ so still more active derivatives can be prepared by using still larger substituents in the meta and para positions. The coefficients of E_{S-2} and E_{S-4} in equation 286 are positive and this, combined with the fact that the larger the group the larger E_S values are in a negative sense, implies that there are obstacles in the C_2 and C_4 space on the enzyme which hinder the entry of large groups in these positions in XXXV. The role of MR for substituents is an ambivalent one. MR may be evaluating dispersion forces binding the inhibitor to the enzyme (198) or it may gauge the volume of the substituent and its ability to distort the conformation of the enzyme so as to preclude interaction with the proper groups on the inhibitor.

Inhibition of guinea pig complement*by benzylpyridinium ions (199-201) and benzamedines (202-206) has been studied by Baker and his collaborators and these results have been analyzed by Hansch *et al* (60,67).

---

* Complement, as described by Hansch and Yoshimoto (60), "consists of 11 distinct proteins which are required for cell lysis brought about via antibodies and complement. The function of the antibody is to identify the invading cell as a foreign organism and activate complement attack which results in cell lysis by means of the proteolytic enzymes." This mechanism operates in the rejection of foreign tissue and organ transplants.

From the complex nature of the benzamidines, XXXVI, it was apparent that dummy variables would have to be employed to encompass the diverse structures present in the 108 benzamidines.



XXXVI

The structural features of substituent X in XXXVI were para meterized by π_1 , σ_1 and MR_1 while π_2, σ_2 and MR_2 were assigned to the Y substituent in the other benzene ring. The indicator variable D_1 characterized the grouping inside the box of XXXVI. For XXXVI compounds with a bridge [$-O(CH_2)_2O-$, $-(CH_2)_4-$, $-O(CH_2)_3O-$, $-O(CH_2)_3-$, $-O(CH_2)_4O-$, $-O(CH_2)_4-$] to a second ring II,* D_1 was set at 1.00. For all other inhibitors D_1 was assigned a value of 0.00. D_2 was introduced to categorize a pyridine ring at the end of the side chain. The connecting side chain was evaluated by π or MR. There was an abnormal exaltation in inhibitory activity when the grouping $\overset{O}{\overset{\|}{NHC}} - Z - C_6H_5$ was attached at C_3 of ring II of XXXVI. When Z was zero, NH, CH_2 , $NHCH_2CH_2$ or CH_2O, D_3 was assigned a value of 1.00. When these groups are at C_2 or C_4 of ring II their contribution was accounted for by π and MR.

---

* D_1 includes the terminal benzene ring, II.

This analysis (60) was approached in the same way as outlined above, and briefly it may be stated that equations 291 and 292 provided the best fit for the data. The correlations from equations 291 and 292 are so similar

$$\log \frac{1}{C} = 0.146\ MR_{1,2} + 1.068\ D_1 + 0.520\ D_2 + 0.429\ D_3 + 2.425$$
$$n = 108, \quad = 0.935, s = 0.258. \ . \ . \qquad 291$$

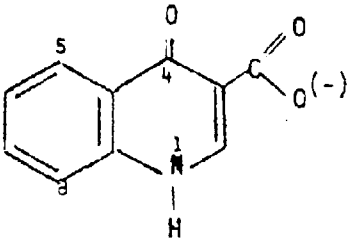
$$\log \frac{1}{C} = 0.211\ \pi_{1,2} + 1.345\ D_1 + 0.620\ D_2 + 0.565\ D_3 + 2.440$$
$$n = 108, r = 0.931, s = 0.267. \ . \ . \qquad 292$$

that it is not possible to say with certainty whether the substituent effect* is hydrophobic in nature or due to the polarizability of the substituents. Hansch and Yoshimoto (60) as well as Coats (207) present circumstantial evidence favouring $MR_{1,2}$ as the descriptor and not π. Expansion of equations 291 and 292 to include $(MR_{1,2})^2$ or $(\pi_{1,2})^2$ terms did not reduce the variance in $\log \frac{1}{C}$ so $\log \frac{1}{C}$ is not a parabolic function of $MR_{1,2}$ or $\pi_{1,2}$. The positive coefficient of $MR_{1,2}$ indicates that the larger the substituents X and Y, the more effective is the inhibitor. Resolution of $MR_{1,2}$ into MR_1 and MR_2 gave an equation in which the coefficients of these two terms were the same sign and about the same magnitude, so the inhibitory effect due to the substituents in each ring is about the same. Hansch and Yoshimoto concluded that MR is not reflecting conformational changes in complement by the substituents but is, more likely, evaluating the binding of the inhibitor to the complement by dispersion forces. Hansch and Yoshimoto then proceeded to analyze the dummy parameters D_1, D_2 and D_3.

The inhibition of complement by 69 benzylpyridinium ions (67) has been treated similarly as has that (74) of chymotrypsin, trypsin, thymidine, phosphorylase, uridine phosphorylase, thymidilate synthetase, cytosine, nucleoside deaminase, malate dehydrogenase, glutamate dehydrogenase, lactate dehydrogenase and glyceraldehyde-phosphate dehydrogenase. The QSAR for the inhibition of malate and glutamate dehydrogenase by 1,4-dihydro-4-quinolone-3-carboxylate ions, XXXVII, are very similar as seen from equations 293 and 294.

---

* The degree of collinearity between π and MR is presented in the correlation matrix in Table III of reference 60.

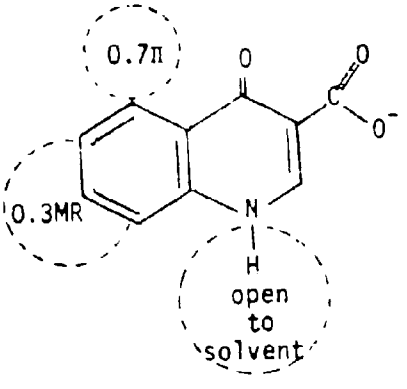


XXXVII

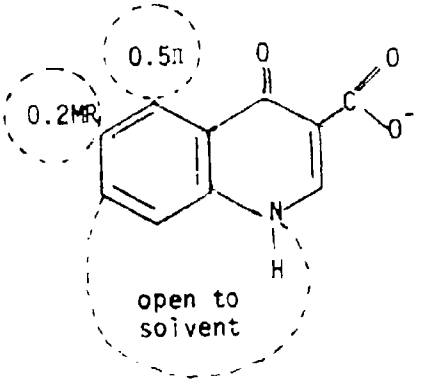
$$\log \frac{1}{C} = 0.699\ \pi_5 + 0.290\ MR_{6,7,8} - 1.121\ I_1 + 3.156$$
$$n = 75,\ r = 0.943,\ s = 0.385 \quad \ldots\ldots\ldots \quad 293$$

$$\log \frac{1}{C} = 0.491\ \pi_5 + 0.233\ MR_6 - 0.553\ I_1 + 3.355$$
$$n = 87,\ r = 0.948,\ s = 0.253 \quad \ldots\ldots\ldots \quad 294$$

From the equations derived for the inhibition of these two enzymes by 1,4-dihydroquinolone-3-carboxylates Yoshimoto and Hansch (74) presented the schematic drawings XXXVIII and XXXIX for enzymic space relative to the 1,4-dihydroquinolone-3-carboxylates. A similar



XXXVIII



XXXIX

diagram was presented (74) for the much more complex process of inhibition of dihydrofolate reductase by diamino-1,3,5-triazines.

Systematic Drug Design

The ideal chemotherapeutic agent is one that is non-toxic to humans but quite toxic to parasitic or malignant cells. A most rational approach to this study is the examination in isolation of enzymic or membrane-controlled processes in both host and parasite. If a chemotherapeutic is known for these processes then systematic drug modification can start with this compound or with one with similar chemical properties and log P value. The systematic modification will involve planned molecular changes that become progressively more extensive. The extensive and classical work of Baker and his collaborators illustrates this approach which he has subdivided into four steps which have been concisely summarized by Hansch (40).

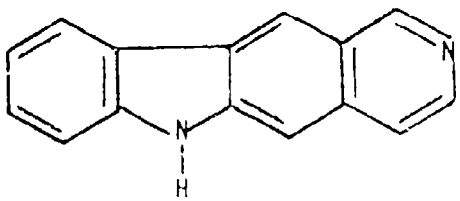
1) "An enzyme must be selected and a reversible inhibitor* found. Modification of its chemically more active groups will lead to suitable reversible inhibitors. Binding points on the reversible inhibitor that complex with the enzyme should be determined. Careful systematic variation of substituents yields the necessary insight.
2) Areas on the inhibitor should be determined in which bulky groups can be placed. This uncovers two types of positions: a) large flexible hydrophobic areas and b) non-contact areas between inhibitor and enzyme.
3) Once the non-contact area is determined, then a group that can form a covalent bond with common enzymic functions should be placed in this area. The length of the side chain by which the function is attached to the parent inhibitor can be varied so that the function can react irreversibly with a group on the enzyme outside the active site.

---

* The most effective inhibitor would be an irreversible inhibitor bound to the enzyme by covalent bonds.

4) After finding the ideal length and flexibility of the side chain which is to react irreversibly, then variations in the active function itself should be investigated in order to find the function with the ideal stereoelectronic specificity, i.e., one whose reactions with the other molecules in the host are minimized."

Careful systematic variation of substituents has been approached in two ways. 1) Topliss (96,97) has developed a systematic approach to the selection of six substituents that will reveal the substituent trend as related to biological activity. 2) Single parameter equations developed from few well selected derivatives will reveal the influence of substituents upon biological activity. The application of the Topliss tree to the study of substituent trends is clearly demonstrated by Hansch's application to the study of ellipticene derivatives based upon the parent compound, XL. (41).



XL

Move 1 in Fig. IX is to introduce a bromine atom at C_5 in XL and determine the biological activity of the 5-bromo-derivative. The 5-bromo-derivative can be more active (+), less active (-) or about the same activity (0) as the parent compound. If the 5-bromo-derivative is more active than XL (right hand side of Fig. IX), then the 5-SCF_3* derivative of XL is prepared and tested. If it is more active than the 5-SCF_3

---

* Both π and σ for SCF_3 are larger in a positive sense than are those for bromine.

THE TOPLISS TREE

XL

Parent Compound

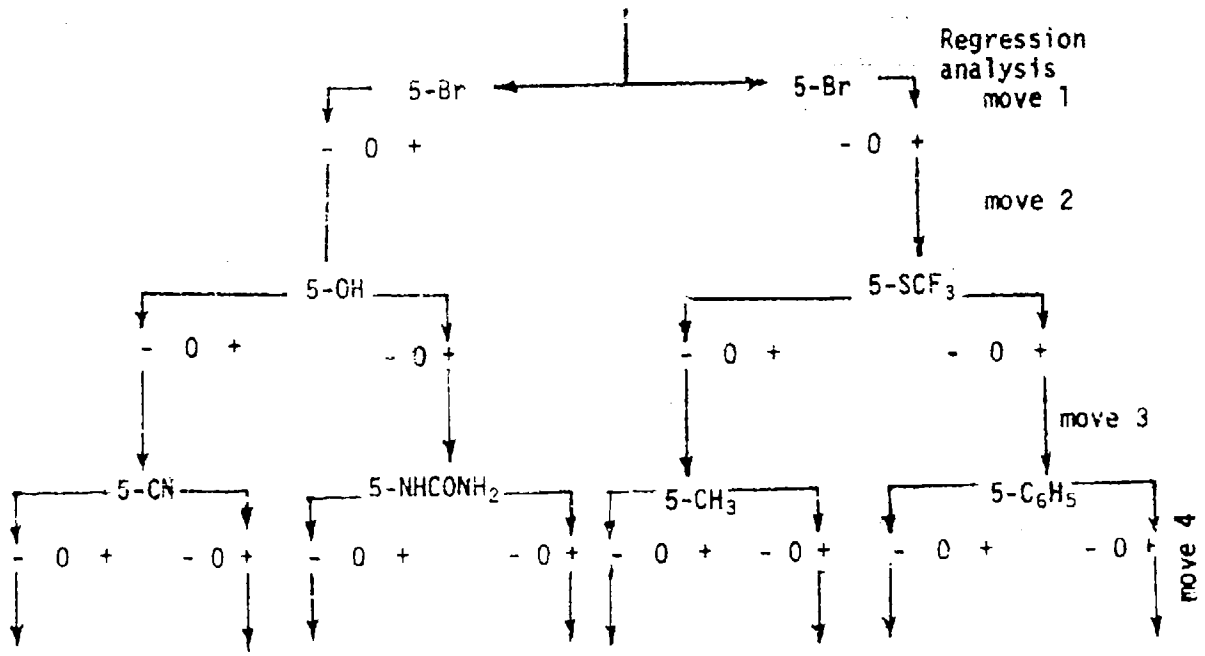


FIG. IX

derivative, movement is to the left and the 5-phenyl derivative is prepared; if less active, then the 5-CH_3 derivative is prepared. Introduction of the phenyl group enhances π but not σ relative to 5-SCF_3. The π and σ constants for 5-CH_3 are lower, in a positive sense, than are those for 5-SCF_3.

If the biological activity of the 5-bromo-derivative is less (-) than that of XL, movement in Fig. IX is to the left and the 5-hydroxy-derivative of XL is prepared and tested. If the 5-hydroxy-derivative is more active (+) then a more water soluble one still, the 5-NH_2CONH-derivative could be prepared and tested. If the result of the 5-OH derivative is (-), it could be that a compound with a positive σ and a negative π value is needed. The 5-CN would meet these requirements nicely.

In the study of the inhibition of complement by benzamidines, Hansch and Yoshimoto (53) recommend the early attempt to establish thermodynamic relationships by development of single parameter equations on a few well chosen compounds. Equations of the form of equations 295-297 could have been developed long before 108 compounds had been synthesized.

$$\log \frac{1}{C} = 1.28\ D_1 + 2.80$$
$$n = 108,\ r = 0.785,\ s = 0.447 \quad (295)$$

$$\log \frac{1}{C} = 1.05\ D_1 + 0.19\ MR_{1,2} + 2.42$$
$$n = 108,\ r = 0.905,\ s = 0.308 \quad (296)$$

$$\log \frac{1}{C} = 0.99\ D_1 + 0.16\ MR_{1,2} + 0.41\ D_3 + 2.48$$
$$n = 108,\ r = 0.927,\ s = 0.273 \quad (297)$$

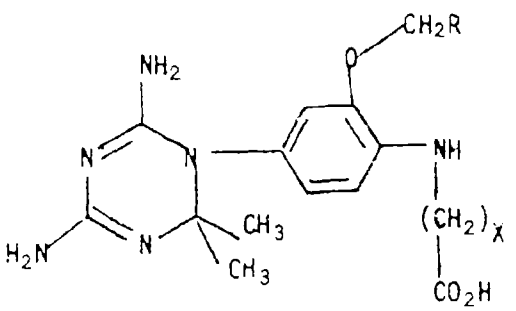
This would have led to better derivatives and eliminated redundancy in the synthesis of derivatives. As a result, equations 291 and 292 would have been obtained in less time and with less work.

Methods for mapping enzymes relative to the inhibitor have been discussed on pages 100 to 110.

The concurrent examination of enzyme or membrane in both host and parasite will ensure attainment of the highest degree of selective toxicity. Realization of a highly active inhibitor on

isolated enzyme does not ensure high activity in *in vivo* tests (see page 84). Competing processes such as metabolism, elimination or anchoring of the inhibitor on the way to the active site may reduce the efficacy of the drug in *in vivo* experiments to almost zero. Log P_0 for another series of drugs acting at the same site in the same animal or that of the same series of drugs acting analogously in another animal (see pages 86 to 90) provides a reasonably safe guideline to follow in selecting the member of the series for the most favourable "random walk" in the *in vivo* experiments.

In general, high rates of metabolism* and elimination are favoured respectively by highly lipophilic and hydrophobic compounds. Equations can be developed for these processes and the sign and magnitude of the coefficients of log P compared with log P_0 for this series of compounds on this *in vivo* biological process (see pages 83 to 90). Employing these methods led Hansch and Silipo (58) to the synthesis of the diamino-1,3,5-triazine, XLI, as a more effective inhibitor of dihydrofolate reductase.



XLI

---

* Metabolites, in general, are quite inactive but compounds like the tetrahydrocannabinols (192) can be encountered where the primary metabolite is the active constituent.

In conclusion, methods have been developed for the calculation of the biological activities of a series of drugs on one organism from their observed activities on another organism (see pages 69 to 73). Preliminary work (84) indicates that this method can be applied to the calculation of biological activities of a series of drugs on man from their observed activities on test animals.

REFERENCES

1. C. Hansch, P.P. Maloney, T. Fujita and R.M. Muir, Nature, 194, 178 (1962).
2. C. Hansch, R.M. Muir, T. Fujita, P.P. Maloney, F. Geiger and M. Streich, J. Amer. Chem. Soc. 85, 2817 (1963).
3. C. Hansch and T. Fujita, J. Amer. Chem. Soc. 86, 1616 (1964).
4. T. Fujita, J. Iwasa and C. Hansch, J. Amer. Chem. Soc. 86, 5175 (1964).
5. C. Hansch, E.W. Deutsch and R.N. Smith, J. Amer. Chem. Soc. 87, 2738 (1965).
6. C. Hansch, K. Kiehs and G.L. Lawrence, J. Amer. Chem. Soc. 87, 5770 (1965).
7. C. Hansch and A.R. Steward, J. Med. Chem. 7, 691 (1964).
8. J. Iwasa, T. Fujita and C. Hansch, J. Med. Chem. 8, 150 (1965).
9. C. Hansch and E.W. Deutsch, J. Med. Chem. 8, 705 (1965).
10. C. Hansch, A.R. Steward and J. Iwasa, J. Med. Chem. 8, 868 (1965).
11. C. Hansch, A.R. Steward and J. Iwasa, Mol. Pharmacol. 1, 87 (1965).
12. C. Hansch, A.R. Steward, J. Iwasa and E.W. Deutsch, Mol. Pharmacol. 1, 205 (1965).
13. K. Kiehs, C Hansch and L. Moore, Biochemistry, 5, 2602 (1966).
14. C. Hansch and E.W. Deutsch, Biochim. Biophys. Acta., 112, 381 (1966).
15. C. Hansch and E.W. Deutsch, Biochim. Biophys. Acta., 126, 117 (1966).
16. C. Hansch and S.M. Anderson, J. Med. Chem. 10, 745 (1967).
17. E. Miller and C. Hansch, J. Pharm. Sci. 56, 92 (1967).
18. C. Hansch, A.R. Steward, S.M. Anderson and D. Bentley, J. Med. Chem. 11, 1 (1968).

19. C. Hansch and S.M. Anderson, J. Org. Chem. 32, 2583 (1967).
20. C. Hansch, J.E. Quinlan and G.L. Lawrence, J. Org. Chem. 33, 347 (1968).
21. C. Hansch, J. Med. Chem. 11, 920 (1968).
22. C. Hansch and E.J. Lien, Biochem. Pharmacol. 17, 709 (1968).
23. C. Hansch, Proc. Int. Pharmacol. Meet. 3rd (1968) 141.
24. C. Hansch, E.J. Lien and F. Helmer, Arch. Biochem. Biophys. 128, 319 (1968).
25. E.J. Lien and C. Hansch, J. Pharm. Sci. 57, 1027 (1968).
26. E. Kutter and C. Hansch, Arch. Biochem. Biophys. 135, 126 (1969).
27. E. Kutter and C. Hansch, J. Med. Chem. 12, 647 (1969).
28. C. Hansch, E. Kutter and A. Leo, J. Med. Chem. 12, 746 (1969).
29. A. Leo, C. Hansch and C. Church, J. Med. Chem. 12, 766 (1969).
30. C. Hansch, J. Org. Chem. 35, 620 (1970).
31. E. Kutter and C. Hansch, Arch. Biochem. Biophys. 135, 126 (1969).
32. C. Hansch and E. Coats, J. Pharm. Sci. 59, 731 (1970).
33. C. Hansch and K.N. VonKaulla, Biochem. Pharmacol. 19, 2193 (1970).
34. C. Hansch, J. Med. Chem. 13, 964 (1970).
35. C. Hansch and E.J. Lien, J. Med. Chem. 14, 653 (1971).
36. A. Leo and C. Hansch, J. Org. Chem. 36, 1539 (1971).
37. C. Hansch, Ann. Rep. Med. Chem. (1966) 347.
38. C. Hansch and W.R. Glave, Molec. Pharmacol. 7, 337 (1971).
39. C. Hansch, J. Schaeffer and R. Kerley, J. Biol. Chem. 247, 4703 (1972).
40. C. Hansch, Annals. NY. Acad. Sci. 186, 235 (1971).
41. C. Hansch, Cancer Chemotherap. Rep. 56, 433 (1972).
42. C. Hansch, J.L. Coubeils and A. Leo, Chim. therap. 6, 427 (1972).
43. C. Hansch, N. Smith, R. Engle and H. Wood, Cancer Chemotherap. Reports 56, 443 (1972).
44. C. Hansch, I.E.P.T. Section V. Chapter 3. Pergamon Press, Oxford, 1973. p. 75.
45. C. Hansch, Advances in Chemistry Series, 114, 20 (1973).
46. C. Hansch and W.J. Dunn, J. Pharm. Sci. 61, 1 (1972).
46a. W.R. Glave and C. Hansch, J. Pharm. Sci. 61, 589 (1972).
47. F. Helmer, K. Kiehs and C. Hansch, Biochem. 7, 2858 (1968).

48. C. Hansch and J.M. Clayton, J. Pharm. Sci. 62, 1 (1973).
49. M.E. Wolff and C. Hansch, Experientia, 29, 1111 (1973).
50. R.N. Smith, C. Hansch and T.P. Poindexter, Physiol. Chem. and Physics 6, 323 (1974).
51. C. Silipo and C. Hansch, Molec. Pharmacol. 10, 954 (1974).
52. W.J. Dunn and C. Hansch, Chem-Biol Interactions, 9, 75 (1974).
53. C. Hansch and M. Yoshimoto, J. Med. Chem. 17, 1160 (1974).
54. C. Hansch, Cancer Chemotherap. Rep. Part 2. 4, 51 (1974).
55. C. Hansch, A. Leo and D.Elkins, J. Chem. Document. 14, 57 (1974).
56. D. Elkins, A. Leo and C. Hansch, J. Chem. Document. 14, 65 (1974).
57. J.A. Montgomery, J. G. Mayo and C. Hansch, J. Med. Chem. 17, 477 (1974).
58. C. Hansch and C. Silipo, J. Med. Chem. 17, 661 (1974).
59. M.E. Wolff and C. Hansch, J. Med. Chem. 17, 898 (1974).
60. C. Hansch and M. Yoshimoto, J. Med. Chem. 17, 1160 (1974).
61. F.R. Quinn, J.S. Driscoll and C. Hansch, J. Med. Chem. 18, 332 (1975).
62. C. Hansch, A. Vittoria, C. Silipo and P.C.Y. Jow, J. Med. Chem. 18, 546 (1975).
63. A. Leo, P.Y.C. Jow, C. Silipo and C. Hansch, J. Med. Chem. 18, 865 (1975).
64. M. Yoshimoto, K.N. Von Kaulla and C. Hansch, J. Med. Chem. 18, 950 (1975).
65. C. Hansch, C. Silipo and E.E. Steller, J. Pharm. Sci., 64 1186 (1975).
66. C. Silipo and C. Hansch, Il. Farmaco, 30, 35 (1975).
67. M. Yoshimoto, C. Hansch, and P.Y.C. Jow, Chem. Pharm. Bull. 23, 437 (1975).
68. C. Hansch, R.N. Smith and R. Engle, Pharmacological Basis of Cancer Chemotherapy, The Williams & Wilkins Co., Baltimore, 1975 p. 756 (215).
69. C. Silipo and C. Hansch, J. Amer. Chem. Soc. 97, 6849 (1975).

70. C. Hansch, J. Med. Chem. 19, 1 (1976).
71. C. Hansch, A. Leo and D. Elkins, J. Chem. Document. 14, 57 (1974).
72. D. Elkins, A. Leo and C. Hansch, J. Chem. Document. 14, 65 (1974).
73. C. Silipo and C. Hansch, J. Med. Chem. 19, 62 (1976).
74. M. Yoshimoto and C. Hansch, J. Med. Chem. 19, 71 (1976).
75. E.J. Lien and G.L. Tong, Cancer Chemother. Rep. 57, 251 (1973).
76. E.J. Lien and C. Hansch, Advances in Chemistry Series, No. 114, 155-182 (1973).
77. E.J. Lien, Proc. 4th Intern. Symp. Med. Chem. The Netherlands, Sept. 1974, pp 319-342.
78. E.J. Lien and C.T. Kong, Pest. Biochem. Physiol. 4, 289 (1974).
79. E.J. Lien, Drug Design. V, Academic Press, New York, 1975. p.81.
80. T. Fujita, Advances in Chemistry Series, 114, 1 (1973).
81. M. Uchida, N. Kurihara, T. Fujita, and M. Nakajima, Pest. Biochem. and Physiol. 4, 260 (1974).
82. T. Fujita, K. Kamoshita, T. Nishioka and N. Nakajima, Agr. Biol. Chem. 38, 1521 (1974).
83. T. Kotani, I. Ichimoto, C. Tatsumi and T. Fujita, Agr. Biol. Chem. 39, 1311 (1975).
84. H.L. Holmes, Structure-Activity Relationships, Defence Research Establishment Suffield, Ralston, Alberta, 1975.
85. A. Spinks, Chem. & Ind. (London) 885 (1973).
86. A. Crum-Brown and T. Fraser, Trans. R. Soc. Edinburgh, 25, 151, 693 (1868-1869).
87. H. Meyer, Arch. exper. Path. Pharmacol. 42, 109 (1899); 46, 338 (1901). See A. Burger Medicinal Chemistry, Interscience Publishers Inc. New York, 1951. p.75.
88. E. Overton, Arch. ges. Physiol. 92, 115 (1902). See A. Burger, Medicinal Chemistry, Interscience Publishers Inc. New York, 1951. p. 75.
89. L.P. Hammett, Chem. Reviews, 35, 125 (1935).
90. L.P. Hammett, Physical Organic Chemistry, McGraw-Hill Book Company, New York, 1940. p. 184.
91. R.W. Taft, Steric Effects in Organic Chemistry, Editor M.S. Newman, John Wiley and Sons, New York, 1956. p. 556.

92. C.Hansch, A. Leo, S.T. Unger, K.H. Kim, D. Nikaitani and E.J. Lien, J. Med. Chem. 16, 1207 (1973).
93. T.C. Bruice, N. Karasch and R.J. Winzler, Arch. Biochem. Biophys. 62, 305 (1956).
94. S.M. Free and J.W. Wilson, J. Med. Chem. 7, 395 (1964).
95. W.P. Purcell, G.E. Bass and J.M. Clayton, Strategy of Drug Design, Wiley-Interscience New York, 1973.
96. J.G. Topliss, J. Med. Chem. 15, 1006 (1972).
97. J.G. Topliss and R.C. Martin, in Drug Design, V.E.J. Ariens, Ed., Academic Press, New York, 1975. p. 1.
98. P.N. Craig, J. Med. Chem. 14, 680, 1251 (1971).
99. C. Hansch, S.H. Unger, and A.B. Forsythe, J. Med. Chem. 16, 1217 (1973).
100. G. Hitchings, Cancer Res. 29, 1895 (1969).
101. B.R. Baker, Design of Active-Site-Directed Irreversible Enzyme Inhibitors, John Wiley and Sons, New York, 1967.
102. J. Drews and F.E. Hahn, Drug Receptor Interactions in Anti-microbial Chemotherapy, Springer-Verlay, Vienna, Austria 1975.
103. H.H. Jaffé, Chem. Reviews, 53, 191 (1953).
104. R.W. Taft, in Steric Effects in Organic Chemistry, John Wiley and Sons, New York, 1956. p. 619.
105. H.C. Brown and Y. Okamoto, J. Amer. Chem. Soc. 80, 4979 (1958).
106. R.W. Taft and I.C. Lewis, J. Amer. Chem. Soc. 80, 2436 (1958).
107. R.W. Taft and I.C. Lewis, J. Amer. Chem. Soc. 81, 5343 (1959).
108. C.G. Swain and E.C. Lupton, J. Amer. Chem. Soc. 90, 4328 (1968).
109. C. Hansch and R. Kerley, Chem & Ind. (London) 294 (1969).
110. T. Yamamoto and T. Otsu, Chem. & Ind. (London) 787 (1967).
111. D.J. Currie, C.E. Lough, R.F. Silver and H.L. Holmes, Can. J. Chem. 44, 1035 (1966).
112. C.E. Lough, R.F. Silver and F.K. McClusky, Can. J. Chem. 46, 1943 (1968).
113. A.D. Delaney, D.J. Currie and H.L. Holmes, Can. J. Chem. 47, 3273 (1969).
114. D.J. Currie and H.L. Holmes, Can. J. Chem. 48, 1340 (1970).
115. C. Hansch, A. Leo and D. Elkins, Chem. Reviews 71, 526 (1971).
116. G.H. Williams, Homolytic Aromatic Substitution, Pergamon Press, New York, 1960.

117. G.H. Williams, Chem. & Ind. (London) 1286 (1961).
118. R.B. Herman, H.W. Culp, R.E. McMahon, and M.M. Marsh, J. Med. Chem. 12, 749 (1969).
119. B. Pressman, D. Swingle, M. Grossberg and A.L. Pauling, J. Amer. Chem. Soc. 66, 1731 (1944).
120. B.R. Baker and H. Shapiro, J. Pharm. Sci. 55, 308 (1966).
121. J.J. Blanksma and D. Hoegen, Rec. Trav. Chim. Pays-Bas, 65, 333 (1946).
122. C.F. Wilkinson, J. Agr. Food Chem. 15, 139 (1967).
123. D.J. Hennessy, J. Agr. Food Chem. 13, 218 (1965).
124. E.R. Garrett, O.K. Wright, G.H. Miller and K.L. Smith, J. Med. Chem. 9, 203 (1966).
125. R.L. Metcalf and T.R. Fukuto, J. Agr. Food Chem. 13, 220 (1965).
126. R.W. Fuller, M.M. Marsh and J. Mills, J. Med. Chem. 11, 397 (1968).
127. A.F. Harms and W.T. Nauta, J. Med. Chem. 2, 57 (1960).
128. C.R. Ensor, D. Russell and G. Chen, J. Pharmacol. Exptl. Therap. 112, 318 (1954).
129. B.R. Baker and G.J. Lourens, J. Med. Chem. 10, 1113 (1967).
130. B.R. Baker and G.J. Lourens, J. Med. Chem. 11, 26 (1968).
131. B.R. Baker and G.J. Lourens, J. Med. Chem. 11 34 (1968).
132. B.R. Baker and G.J. Lourens, J. Med. Chem. 11, 38 (1968).
133. B.R. Baker, J. Med. Chem. 11, 483 (1968).
134. B.R. Baker and G.J. Lourens, J. Med. Chem. 11, 666 (1968).
135. B.R. Baker and G.J. Lourens, J. Med. Chem. 11, 672 (1968).
136. B.R. Baker and G.J. Lourens, J. Med. Chem. 11, 677 (1968).
137. B.R. Baker and G.J. Lourens, J. Med. Chem. 12, 92 (1969).
138. B.R. Baker and G.J. Lourens, J. Med. Chem. 12, 95 (1969).
139. B.R. Baker and G.J. Lourens, J. Med. Chem. 12, 101 (1969).
140. B.R. Baker and M.A. Johnson, J. Med. Chem. 11, 486 (1968).
141. W.T. Ashton, L.L. Kirk and B.R. Baker, J. Med. Chem. 16, 453 (1973).
142. B.R. Baker, G.J. Lourens, R.B. Meyer and N.M.J.Vermeulen, J. Med. Chem. 12, 67 (1969).

143. B.R. Baker and E.E. Janson, J. Med. Chem. 12, 672 (1969).
144. B.R. Baker and W.T. Ashton, J. Med. Chem. 12, 894 (1969).
145. B.R. Baker and W.T. Ashton, J. Med. Chem. 13, 1149 (1970).
146. B.R. Baker and W.T. Ashton, J. Med. Chem. 13, 1161 (1970).
147. B.R. Baker and W.T. Ashton, J. Med. Chem. 13, 1165 (1970).
148. B.R. Baker and W.T. Ashton, J. Med. Chem. 15, 945 (1972).
149. B.R. Baker and W.T. Ashton, J. Med. Chem. 16, 209 (1973).
150. B.R. Baker, E.E. Janson and N.M.J. Vermeulen, J. Med. Chem. 12, 898 (1969).
151. B.R. Baker, N.M.J. Vermeulen, W.T. Ashton and A.J. Ryan, J. Med. Chem. 13, 1130 (1970).
152. A.J. Ryan, N.M.J. Vermeulen,and B.R. Baker, J. Med.Chem. 13 1140 (1970).
153. B.R. Baker and N.M.J. Vermeulen, J. Med. Chem. 13, 1143 (1970).
154. B.R. Baker and N.M.J. Vermeulen, J. Med. Chem. 13, 1154 (1970).
155. E.W. Maynert and H.B. Van Dyke,Pharmacol. Rev. 1, 217 (1949).
156. A. Dorfman and L.R. Goldbaum, J. Pharmacol. Exptl. Therap. 90, 330 (1947).
157. B.B. Brodie, J.R. Gillette and B.N. LaDu, Ann. Rev. Biochem. 27, 427 (1958).
158. Y.C. Martin and C. Hansch, J. Med. Chem. 14, 777 (1971).
159. W.H. Schmidt and A.J. Moyer, J. Bacteriol. 47, 199 (1944).
160. F. Fichter and W. Dietrich, Helv. Chim. Acta, 7, 137 (1924).
161. W.W. Westerfield and C. Lowe, J. Biol. Chem. 145, 463 (1942).
162. K. Kratzl, J. Kratzl and P. Claus, Advances in Chem. Ser. 59, 157 (1966).
163. H. Booth and B.C. Saunders, J. Chem. Soc. 940 (1956).
164. H. Richtenzain, Chem. Ber. 82, 447 (1949).
165. R.H.F. Manske and H.L. Holmes, The Alkaloids, Vol. II, Academic Press Inc. New York, 1952. pp 1-217.
166. K.W. Bentley and S.F. Dyke, J. Chem. Soc. 2574 (1959).
167. T.J. Wallace, J.M. Miller, H. Probner and A. Schriesheim, Proc. Chem. Soc. 243 (1962).
168. L.F. Fieser, J. Amer. Chem. Soc. 52, 5204 (1930).
169. D.E. Pennington and D.M. Ritter, J. Amer. Chem. Soc. 69 46, 187 (1947).

170. W.A. Waters, J. Chem. Soc. Ser. B, 2026 (1971).
171. A.A. Yasnikov and E.M. Gaivoronskaya, Ukrain. Khim. Zhur. 27, 506 (1961); C.A. 56, 7205 (1962).
172. C.E. Williamson and B. Witten, U.S. Clearing House Fed. Sci. Tech. Inform. AD 1970 No. 723408; C.A. 75, 116969 (1971).
173. O. Schales and H.A. Graefe, J. Amer. Chem. Soc. 74, 4486 (1952).
174. H. Kreuger, N.B. Eddy and M. Sumwalt, U.S. Public Health Suppl. No. 165, Vol. II, 1943. p. 943.
175. E.L. May and L.J. Sargent, Medicinal Chemistry Monograph Series, Vol. 5, Academic Press, New York, 1965. pp 131, 141-143.
176. N.B. Eddy and D. Leimbach, J. Pharmacol. Exptl. Therap. 107, 385 (1953).
177. K.W. Bentley, The Chemistry of Morphine Alkaloids, Clarendon Press, Oxford, 1954.
178. J.C. McGowan, P.W. Brian and H.G. Hemming, Ann. Applied Biol. 35, 25 (1948).
179. T. Ban and T. Fujita, J. Med. Chem. 12, 353 (1969).
180. J.G. Beasley and W.P. Purcell, Biochim. Biophys. Acta. 178, 175 (1969).
181. A.H. Solway, B. Whitman and J.R. Messer, J. Pharmacol. Exptl. Therap. 129, 310 (1960): P.N. Craig, J. Med. Chem. 15, 144 (1972).
182. A.C. Cope and E.M. Hancock, J. Amer. Chem. Soc. 61, 353 (1939): J.M. Clayton and W.P. Purcell, J. Med. Chem. 12, 1087 (1969).
183. H.A. Shonle and A. Moment, J. Amer. Chem. Soc. 45, 243 (1923).
184. D.L. Tabern and E.H. Volwiler, J. Amer. Chem. Soc. 56, 1139 (1934).
185. W.J. Doran and H.A. Shonle, J. Amer. Chem. Soc. 59, 1625 (1937).
186. E.H. Volwiler, J. Amer. Chem. Soc. 47, 2236 (1925).
187. A.C. Cope, P. Kovacic and M. Burg, J. Amer. Chem. Soc. 71, 3658 (1949).
188. A.C. Cope and E.M. Hancock, J. Amer. Chem. Soc. 61, 776 (1939).
189. A.C. Cope, W.H. Hartung, E.M. Hancock and F.S. Crossley, J. Amer. Chem. Soc. 62, 1199 (1940).
190. E.J. Lien, C. Hansch and S.M. Anderson, J. Med. Chem. 11, 430 (1968).

191. C. Hansch, E.W. Deutsch and R.N. Smith, J. Amer. Chem. Soc. 87, 2738 (1965).
192. R.S. Wilson and E.L. May, J. Med. Chem. 18, 700 (1975).
193. K.W. Bentley and D.G. Hardy, J. Amer. Chem. Soc. 89, 3281 (1967).
194. B.R. Baker and H. Shapiro, J. Pharm. Sci., 55, 308 (1966).
195. T.E. Gram and J.R. Fouts, J. Pharmacol. Exptl. Therap. 152, 363 (1966).
196. T.E. Gram, L.A. Rogers and J.R. Fouts, J. Pharmacol. Exptl. Therap. 155, 479 (1967).
197. B.R. Baker and W.F. Wood, J. Med. Chem. 10, 1101 (1966).
198. D. Agin, L. Hersh and D. Holtzman, Proc. Nat. Acad. U.S. 53, 932 (1965): B.R. Baker and W.F. Wood, J. Med. Chem. II 644 (1967).
199. B.R. Baker and J.A. Hurlbut, J. Med. Chem. 12, 677 (1969).
200. B.R. Baker and J.A. Hurlbut, J. Med. Chem. 12, 902 (1969).
201. B.R. Baker and M.H. Doll, J. Med. Chem. 14, 793 (1971).
202. B.R. Baker and E.H. Erickson, J. Med. Chem. 12, 408 (1969).
203. B.R. Baker and M. Cory, J. Med. Chem. 12, 1049 (1969).
204. B.R. Baker and M. Cory, J. Med. Chem. 12, 1053 (1969).
205. B.R. Baker and M. Cory, J. Med. Chem. 14, 119 (1971).
206. B.R. Baker and M. Cory, J. Med. Chem. 14, 805 (1971).
207. E.A. Coats, J. Med. Chem. 16, 1102 (1973).

SUBJECT INDEX*

* For abbreviations and symbols see pages 3 - 5.

A

B

UNCLASSIFIED

Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall document is classified)

1. ORIGINATING ACTIVITY

DEFENCE RESEARCH ESTABLISHMENT SUFFIELD

2a. DOCUMENT SECURITY CLASSIFICATION

UNCLASSIFIED

2b. GROUP

3 DOCUMENT TITLE

DEVELOPMENTS IN QUANTITATIVE STRUCTURE-ACTIVITY RELATIONSHIPS (QSAR) A REVIEW

4 DESCRIPTIVE NOTES (Type of report and inclusive dates) Report

5 AUTHOR(S) (Last name, first name, middle initial)

Holmes, H.L.

6. DOCUMENT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
July 1976	134	207

8a. PROJECT OR GRANT NO.	9a. ORIGINATOR'S DOCUMENT NUMBER(S)
20-03-05	SUFFIELD REPORT NO. 282
8b. CONTRACT NO.	9b. OTHER DOCUMENT NO.(S) (Any other numbers that may be assigned this document)

10. DISTRIBUTION STATEMENT

[redacted]

11. SUPPLEMENTARY NOTES	12. SPONSORING ACTIVITY

13. ABSTRACT

Three different approaches have been developed by groups headed by Hansch, Holmes and by Free and Wilson to the problem of quantitative structure-activity relationships (QSAR) of drugs to biological processes and these have been reviewed and the strengths and weaknesses of each method annotated.....

-1473-

DD 1473
78-078

KEY WORDS

Structure-Activity Relationships
QSAR
Drugs - design
Biological activity

INSTRUCTIONS

1. ORIGINATING ACTIVITY: Enter the name and address of the organization issuing the document.

2a. DOCUMENT SECURITY CLASSIFICATION: Enter the overall security classification of the document including special warning terms whenever applicable.

2b. GROUP: Enter security reclassification group number. The three groups are defined in Appendix 'M' of the DRB Security Regulations.

3. DOCUMENT TITLE: Enter the complete document title in all capital letters. Titles in all cases should be unclassified. If a sufficiently descriptive title cannot be selected without classification, show title classification with the usual one-capital-letter abbreviation in parentheses immediately following the title.

4. DESCRIPTIVE NOTES: Enter the category of document, e.g. technical report, technical note or technical letter. If appropriate, enter the type of document, e.g. interim, progress, summary, annual or final. Give the inclusive dates when a specific reporting period is covered.

5. AUTHOR(S): Enter the name(s) of author(s) as shown on or in the document. Enter last name, first name, middle initial. If military, show rank. The name of the principal author is an absolute minimum requirement.

6. DOCUMENT DATE: Enter the date (month, year) of Establishment approval for publication of the document.

7a. TOTAL NUMBER OF PAGES: The total page count should follow normal pagination procedures, i.e., enter the number of pages containing information.

7b. NUMBER OF REFERENCES: Enter the total number of references cited in the document.

8a. PROJECT OR GRANT NUMBER: If appropriate, enter the applicable research and development project or grant number under which the document was written.

8b. CONTRACT NUMBER: If appropriate, enter the applicable number under which the document was written.

9a. ORIGINATOR'S DOCUMENT NUMBER(S): Enter the official document number by which the document will be identified and controlled by the originating activity. This number must be unique to this document.

9b. OTHER DOCUMENT NUMBER(S): If the document has been assigned any other document numbers (either by the originator or by the sponsor), also enter this number(s).

10. DISTRIBUTION STATEMENT: Enter any limitations on further dissemination of the document, other than those imposed by security classification, using standard statements such as:

(1) "Qualified requesters may obtain copies of this document from their defence documentation center."

(2) "Announcement and dissemination of this document is not authorized without prior approval from originating activity."

11. SUPPLEMENTARY NOTES: Use for additional explanatory notes.

12. SPONSORING ACTIVITY: Enter the name of the departmental project office or laboratory sponsoring the research and development. Include address.

13. ABSTRACT: Enter an abstract giving a brief and factual summary of the document, even though it may also appear elsewhere in the body of the document itself. It is highly desirable that the abstract of classified documents be unclassified. Each paragraph of the abstract shall end with an indication of the security classification of the information in the paragraph (unless the document itself is unclassified) represented as (TS), (S), (C), (R), or (U).

The length of the abstract should be limited to 20 single-spaced standard typewritten lines; 7½ inches long.

14. KEY WORDS: Key words are technically meaningful terms or short phrases that characterize a document and could be helpful in cataloging the document. Key words should be selected so that no security classification is required. Identifiers, such as equipment model designation, trade name, military project code name, geographic location, may be used as key words but will be followed by an indication of technical context.